

Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model

Run on: March 25, 2003, 08:17:33 ; Search time 26.1818 Seconds
 (without alignments)
 45.805 Million cell updates/sec

Title: US-09-646-532B-3

Perfect score: 46 Sequence: 1 NFRYTNFAX 9

Scoring table: BLOSUM62 Gapop 10.0 , Gapext: 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : A_Genesed_101002:*

1: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1980.DAT: *
 2: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1981.DAT: *
 3: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1982.DAT: *
 4: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1983.DAT: *
 5: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1984.DAT: *
 6: /SIDS1/gcadata/geneseqp-emb1/AA1985.DAT: *
 7: /SIDS1/gcadata/geneseqp-emb1/AA1986.DAT: *
 8: /SIDS1/gcadata/geneseqp-emb1/AA1987.DAT: *
 9: /SIDS1/gcadata/geneseqp-emb1/AA1988.DAT: *
 10: /SIDS1/gcadata/geneseqp-emb1/AA1989.DAT: *
 11: /SIDS1/gcadata/geneseqp-emb1/AA1990.DAT: *
 12: /SIDS1/gcadata/geneseqp-emb1/AA1991.DAT: *
 13: /SIDS1/gcadata/geneseqp-emb1/AA1992.DAT: *
 14: /SIDS1/gcadata/geneseqp-emb1/AA1993.DAT: *
 15: /SIDS1/gcadata/geneseqp-emb1/AA1994.DAT: *
 16: /SIDS1/gcadata/geneseqp-emb1/AA1995.DAT: *
 17: /SIDS1/gcadata/geneseqp-emb1/AA1996.DAT: *
 18: /SIDS1/gcadata/geneseqp-emb1/AA1997.DAT: *
 19: /SIDS1/gcadata/geneseqp-emb1/AA1998.DAT: *
 20: /SIDS1/gcadata/geneseqp-emb1/AA1999.DAT: *
 21: /SIDS1/gcadata/geneseqp-emb1/AA2000.DAT: *
 22: /SIDS1/gcadata/geneseqp-emb1/AA2001.DAT: *
 23: /SIDS1/gcadata/geneseqp-emb1/AA2002.DAT: *

Pred. No. 1 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	45	97.8	9	20	AYV42755	Wheat amyloplast A
2	45	97.8	9	20	AYV3933	ADP glucose transp
3	38	82.6	28	12	AAR14566	N-terminal fragmen
4	34	73.9	316	22	ABG24752	Novel human digno
5	33	71.7	436	22	AAU42256	Propionibacterium
6	33	71.7	1180	22	ABG22442	Novel human digno
7	32	69.6	76	22	AAU33056	Novel human secret
8	32	69.6	209	22	ABG84920	Shrimp white spot
9	32	69.6	245	21	AAB07810	A galactanase of B
10	32	69.6	288	21	AAG57319	Arabidopsis thalia

ALIGNMENTS

RESULT 1 ID AY42755		standard; peptide: 9 AA.	
XX		AY42755;	
AC		AY42755;	
XX			
DT		20-DEC-1999 (first entry)	
XX			
DE		wheat amyloplast ADP-glucose transporter peptide #3.	
XX			
KW		Starch biosynthesis; amyloplast; ADP-glucose; transport; import; plant; amylopectin; amylose; branching; chemical structure; transgenic; plant; optimisation; industrial applications.	
KW			
OS		Triticum aestivum.	
XX			
FH		Key' location/Qualifiers	
FT		Misc-difference ⁹ /label= Xaa /note= "Xaa = any amino acid"	
FT			
XX			
PN		W09947682-A1.	
XX			
PD		23-SEP-1999.	
XX			
PF		19-MAR-1999; 99WO-GB00728.	
PR		20-MAR-1998; 98GB-0005939.	
XX			
PA		(UWMA-) UNIV VICTORIA MANCHESTER.	
XX			
PT		Eines MJ, Tetlow IJ, Bowsher CG;	
XX			

DR WPI; 1999-590977/50.
 XX PT New transporter complex protein useful for modulating starch content in
 PT plants, especially useful in food production -
 XX PS Claim 1; Page 3; 28pp; English.

XX This sequence represents a wheat amyloplast ADP-glucose transporter
 CC peptide, #2. The wheat amyloplast ADP-glucose transporter is associated
 CC with the amyloplast membrane and comprises at least two proteins; this
 CC sequence, along with peptides #1 (AY42753) and #2 (AY42754) are
 CC components of one of these proteins. Peptides #4/#7 (AY42756-#42759)
 CC represent sequences within the second protein (AY42760). The sugar
 CC nucleotide ADP-glucose is the immediate substrate for starch synthesis,
 CC which occurs in the amyloplast; however, ADP-glucose is mainly
 CC synthesised outside the amyloplast in the cytoplasm. The ADP-glucose
 CC transporter is responsible for the import of ADP-glucose into the
 CC amyloplast and therefore plays a pivotal role in the regulation of starch
 CC synthesis. The transporter not only influences starch yield, but also
 CC quality as the starch synthases involved in amylose and amylopectin
 CC syntheses have different affinities for ADP-glucose. Variations in the
 CC chemical structure of starch are determined by the ratio of amylose to
 CC amylopectin, and by the degree of branching in amylopectin in the starch
 CC polymer. These variations can significantly alter the properties of
 CC starch. The ADP-glucose transporter complex is useful for generating
 CC transgenic plants in which the starch quality can be optimised for
 CC industrial applications in which starch is used. For example, transgenic
 CC plants which have an increased amylose content in starch are useful for
 CC production of starch with increased viscosity and gel strength, which
 CC prevents baked food going stale so quickly. Conversely, an increased
 CC amylopectin content in the starch produces waxy starch useful as
 CC thickening agents in food and coatings.

XX Sequence 9 AA;

Query Match 97.8%; Score 45; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTNAFA 8
 ID 1|||||||
 Db 1 NFRYTNAFA 8

RESULT 2

AY39333 AAY39333 standard; peptide; 9 AA.

ID XX AC AY39333;
 DT XX DE 01-DEC-1999 (first entry)

DE ADP glucose transporter peptide 3.

XX KW ADP glucose transporter; transform plant cell; wheat; starch production;
 KW waxy starch; thickening agent; food; coating; increased viscosity; stale;
 KW gel strength; baked food.

OS OS Triticum aestivum.

Key FH KW glucose transporter; transform plant cell; wheat; starch production;
 FT waxy starch; thickening agent; food; coating; increased viscosity; stale;
 FT gel strength; baked food.

Misc-difference FT Location/Qualifiers
 /note= "Any amino acid"

PA WO9947681-A1.
 PD 23-SEP-1999.
 PN 19-MAR-1999; 99W0-GB00727.
 PR 20-MAR-1998; 98GB-0005939.
 PA (UIMA-) UNIV VICTORIA MANCHESTER.

XX PI Emes MJ, Tetlow IJ, Bowsher CG;
 XX DR WPI; 1999-571441/48.

XX PT ADP glucose transporter protein used for modifying plant starch
 production -
 XX PS Claim 1; Page 15; 26pp; English.

XX Peptides AAY39331-Y39337 are fragments of an ADP glucose transporter
 CC protein. The protein contains at least one of the peptide sequences, and
 CC is capable of ADP glucose transport. A DNA molecule encoding an ADP
 CC glucose transporter protein can be used to transform plant cells. The
 CC ADP glucose transporter protein can be used to regulate starch
 CC production from a plant. The plants can then be used to produce waxy
 CC starches that can be used as thickening agents in food and coatings.
 CC Alternatively the plants can be used to produce a starch with increased
 CC viscosity and gel strength, which can be incorporated in baked food which
 CC stays fresh for a longer length of time.

XX SQ Sequence 9 AA;

Query Match 97.8%; Score 45; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTNAFA 8
 ID 1 NFRYTNAFA 8

RESULT 3

AARI4366 AARI4366 standard; Peptide; 28 AA.

XX AC AARI4366;
 DT XX DE 30-JAN-1992 (first entry)

XX N-terminal fragment of CTPA 28A32 36K protein.

XX KW Tumour cell; epitope; colon; cancer; monoclonal antibodies; vaccine;
 KW Antigen; carcinoma.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Region 1
 FT /label= unknown
 FT Region 26
 FT /label= unknown

PN WO9116629-A.
 PD 31-OCT-1991.
 PF 10-APR-1991; 91WO-U002459.

PR 12-APR-1990; 90US-0508373.

XX PA (ALKU) AKZO NV.
 XX PI Pomato N, BOS ES, Ransom JH, Hanna MG;
 DR WPI; 1991-339972/46.

XX PT New human MAB 28A32 - used to diagnose and monitor colon cancer
 PT and to prepare vaccines.
 XX PR Claim 1; Fig 7; 34pp; English.

XX CC The sequence was obtd. by Edman degradation and represents the

most predominant and reproducible isolatable fragment of the 36K protein. The protein is one of four (50K, 46K, 36K, and 32K) associated with the CTAA 28A32 antigen. The 36K protein is associated with membranes; the rest are found in the cytoplasm of cells. The antigen is associated with colon tumours and monoclonal antibody 28A32 can be used to monitor and detect the presence of tumour cells. Vaccines can also be prepd. from the proteins. See also AARI4364-67.

SQ Sequence 28 AA;
 Query Match 82.6%; Score 38; DB 12; Length 28;
 Best Local Similarity 87.5%; Pred. No. 2.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NFRYTNEA 8
 Db 15 NFDYTNEA 22

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the Printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 316 AA;
 Query Match 73.9%; Score 34; DB 22; Length 316;
 Best Local Similarity 75.0%; Pred. No. 1 4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NFRYTNEA 8
 Db 195 NFRYKDFA 202

RESULT 4
 ABG24752 standard; Protein: 316 AA.
 XX ABG24752;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #24743.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS88939.

RESULT 5
 AAU42256 standard; Protein: 436 AA.
 XX AAU42256
 XX AC AAU42256;
 XX DT 27-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #3152.
 XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX PN WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US12865.
 XX PR 21-APR-2000; 2000US-199047P.
 XX PR 03-JUN-2000; 2000US-208841P.
 XX PR 07-JUL-2000; 2000US-216747P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Sheik YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX DR WPI; 2001-616774/71.
 XX DR N-PSDB; AAS59516.

XX PT propionibacterium acnes polypeptides and nucleic acids useful for
 XX PT vaccinating against and diagnosing infections, especially useful for
 XX PT treating acne vulgaris -
 XX PS Example 1; SEQ ID No 3451; 1069pp; English.
 XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the protein of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to

CC	downregulate expression and activity of <i>P. acnes</i> polypeptides and therefore treat <i>P. acnes</i> infections. The antibodies may also be used as diagnostic agents for determining <i>P. acnes</i> presence, for example, by enzyme linked immunosorbent assay (ELISA).
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences .
CC	SQ Sequence 436 AA;
XX	Query Match 71.7%; Score 33; DB 22; Length 436;
Best Local Matches 85.7%; Pred. No. 3e+02;	Mismatches 0; Indels 1; Gaps 0;
QY 2 FRYNFA 8	Db 369 FRYNFA 375
XX	Sequence 1180 AA;
RESUL ^T 6	Query Match 71.7%; Score 33; DB 22; Length 1180;
ID ABG22442	Best Local Matches 71.4%; Pred. No. 7.8e-02; Mismatches 1; Indels 0; Gaps 0;
ID ABG22442 standard; Protein; 1180 AA.	
XX	
AC ABG22442;	
XX	
DT 18-FEB-2002 (first entry)	
DE Novel human diagnostic protein #22433.	
XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.	
KW KW Homo sapiens.	
OS OS Homo sapiens.	
PN WO200175067-A2.	
XX	
PD 11-OCT-2001.	
XX	
PF 30-MAR-2001; 2001WO-US08631.	
XX	
PR 31-MAR-2000; 2000US-0540217.	
PR 23-AUG-2000; 2000US-0649167.	
XX	
PA (HYSE-) HYSEQ INC.	
XX	
PI Drmanac RT, Liu C, Tang YT;	
XX	
DR WPI; 2001-639362/73.	
XX	
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity	
PT	
PS Claim 20; SEQ ID NO 52801; 103PP; English.	
XX	
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (I) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging or sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and	
CC amino acid sequences AB000010-AB310377 represent novel human diagnostic amino acid sequences of the invention.	
CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences .	
CC	SQ Sequence 1180 AA;
RESUL ^T 7	Query Match 71.7%; Score 33; DB 22; Length 1180;
ID AAU33056	Best Local Matches 71.4%; Pred. No. 7.8e-02; Mismatches 1; Indels 0; Gaps 0;
ID AAU33056 standard; Protein; 76 AA.	
XX	
AC AAU33056;	
XX	
DT 18-DEC-2001 (first entry)	
DE Novel human secreted protein #3547.	
XX Human; vaccination; gene therapy; nutritional supplement; stem cell proliferation; haemopoiesis; nerve tissue regeneration; immune suppression; immune stimulation; anti-inflammatory; leukaemia.	
KW KW Homo sapiens.	
XX OS Homo sapiens.	
PN WO200179449-A2.	
XX	
PD 25-OCT-2001.	
XX	
PF 16-APR-2001; 2001WO-US08656.	
XX	
PR 18-APR-2000; 2000US-0552029.	
PR 26-JAN-2001; 2001US-0770160.	
XX	
PA (HYSE-) HYSEQ INC.	
XX	
PI Tang YT, Liu C, Drmanac RT;	
XX	
DR WPI; 2001-611725/70.	
XX	
PT Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy	
XX	
PS Claim 20; Page 702; 765PP; English.	
XX	
The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haemopoiesis; and in bone carriage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AU29510-AU33304 represent the amino acid sequences of novel human secreted proteins of the invention.	
CC	SQ Sequence 76 AA;

	Query Match	69.6%	Score	32;	DB	22;	Length	76;	XX
	Best Local Similarity	83.3%	Pred.	No.	82;		Mismatches	0;	AC
	Matches	5;	Conservative	1;	Mismatches	1;	Indels	0;	AAB07810;
QY	2	FRYTNF	7	DE	14-NOV-2000	(first entry)			XX
Db	57	FRYTNF	62	XX					
	RESULT	8							
	AAG84920								
ID	AAG84920	standard; Protein; 209 AA.							
XX									
AC	AAG84920;								
XX									
DT	11-SEP-2001	(first entry)							
XX									
DE	Shrimp white spot; Bacilliform virus (WSBV) protein 11.								
XX									
KW	Shrimp white spot; Bacilliform virus; WSBV; diagnosis; viral infection;								
KW	antiviral agent; gene expression; antisense construct;								
KW	transgenic viral resistant shrimp.								
XX									
OS	white spot syndrome virus.								
XX									
PN	WO200138351-A2.								
XX									
PD	31-MAY-2001.								
XX									
PF	08-NOV-2000; 2000WO-US288888.								
XX									
PR	24-NOV-1999; 99CN-0124717.								
XX									
PA	(PENY-) PE CORP NY.								
PA	(THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.								
PA	(SINO-) SINGONONMAX CO LTD.								
XX									
PT	Novel methods for modifying animal feed using galactanase and novel								
PT	galactanase enzymes useful for modifying animal feed.								
XX									
PS	Claim 38; Page 64-65; 77pp; English.								
XX									
CC	The present sequence represents a <i>Bacillus galactanase enzymes</i> . The								
CC	enzyme comprises at least one of the consensus sequences AAB07102-07,								
CC	and is used in the method of the invention. The specification								
CC	describes a method for modifying animal feed using galactanase.								
CC	Galactanase enzymes degrade galactans. The galactanases of								
CC	the invention are glycosyl hydrolases. The galactanase enzymes are								
CC	useful for the modification of animal feed and in the textile,								
CC	detergent, wine and juice and cellulose processing industries.								
XX									
SQ	Sequence 245 AA;								
	Query Match	69.6%	Score	32;	DB	21;	Length	245;	
	Best Local Similarity	71.4%	Pred.	No.	2.6e+02;	1;	Mismatches	0;	
	Matches	5;	Conservative	1;	Indels	0;	Gaps	0;	
QY	1	FRYTNF	7	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 73849.				
Db	47	NFYSNF	53	XX	Protein identification; signal transduction pathway; metabolic pathway;				
				XX	hybridization assay; genetic mapping; gene expression control; promoter;				
				XX	termination sequence.				
				OS	Arabidopsis thaliana.				
				XX					
				PN	EP033405-A2.				
				PD	06-SEP-2000.				
RESULT	9								
ID	AAB07810	AAB07810 standard; Protein; 245 AA.							

XX	PR	08-JUL-1999;	99US-0142803.
PF	PR	09-JUL-1999;	99US-0142270.
XX	PR	12-JUL-1999;	99US-0142277.
PR	PR	13-JUL-1999;	99US-0143542.
PR	PR	14-JUL-1999;	99US-014324.
PR	PR	15-JUL-1999;	99US-0144005.
PR	PR	16-JUL-1999;	99US-0144085.
PR	PR	16-JUL-1999;	99US-0144086.
PR	PR	17-JUL-1999;	99US-0144325.
PR	PR	19-JUL-1999;	99US-0144331.
PR	PR	19-JUL-1999;	99US-0144332.
PR	PR	19-JUL-1999;	99US-0144333.
PR	PR	19-JUL-1999;	99US-0144334.
PR	PR	19-JUL-1999;	99US-0144335.
PR	PR	21-JUL-1999;	99US-0144336.
PR	PR	20-JUL-1999;	99US-0144632.
PR	PR	20-JUL-1999;	99US-0144884.
PR	PR	21-JUL-1999;	99US-0144814.
PR	PR	21-JUL-1999;	99US-0145086.
PR	PR	21-JUL-1999;	99US-0145088.
PR	PR	22-JUL-1999;	99US-0145087.
PR	PR	22-JUL-1999;	99US-0145776.
PR	PR	22-JUL-1999;	99US-0145913.
PR	PR	22-JUL-1999;	99US-0145192.
PR	PR	23-JUL-1999;	99US-0145145.
PR	PR	23-JUL-1999;	99US-0145218.
PR	PR	23-JUL-1999;	99US-0145324.
PR	PR	26-JUL-1999;	99US-0145324.
PR	PR	02-AUG-1999;	99US-0146388.
PR	PR	02-AUG-1999;	99US-0146389.
PR	PR	03-AUG-1999;	99US-0147038.
PR	PR	03-AUG-1999;	99US-0147919.
PR	PR	04-AUG-1999;	99US-0147302.
PR	PR	04-AUG-1999;	99US-0146386.
PR	PR	05-AUG-1999;	99US-0145324.
PR	PR	05-AUG-1999;	99US-0145324.
PR	PR	06-AUG-1999;	99US-0147260.
PR	PR	06-AUG-1999;	99US-0147303.
PR	PR	06-AUG-1999;	99US-0147038.
PR	PR	07-AUG-1999;	99US-0147416.
PR	PR	07-AUG-1999;	99US-0147493.
PR	PR	09-AUG-1999;	99US-0147335.
PR	PR	10-AUG-1999;	99US-0148171.
PR	PR	11-AUG-1999;	99US-0147192.
PR	PR	12-AUG-1999;	99US-0147260.
PR	PR	13-AUG-1999;	99US-0148565.
PR	PR	13-AUG-1999;	99US-0147926.
PR	PR	16-AUG-1999;	99US-0148084.
PR	PR	17-AUG-1999;	99US-0149175.
PR	PR	18-AUG-1999;	99US-0149226.
PR	PR	20-AUG-1999;	99US-0149226.
PR	PR	20-AUG-1999;	99US-0149723.
PR	PR	20-AUG-1999;	99US-0149929.
PR	PR	23-AUG-1999;	99US-0149002.
PR	PR	23-AUG-1999;	99US-014930.
PR	PR	25-AUG-1999;	99US-0150666.
PR	PR	26-AUG-1999;	99US-0150884.
PR	PR	27-AUG-1999;	99US-0152363.
PR	PR	27-AUG-1999;	99US-0151666.
PR	PR	27-AUG-1999;	99US-0153070.
PR	PR	30-AUG-1999;	99US-0151080.
PR	PR	31-AUG-1999;	99US-0151303.
PR	PR	16-SEP-1999;	99US-0151438.
PR	PR	16-SEP-1999;	99US-0151330.
PR	PR	20-SEP-1999;	99US-0154179.
PR	PR	22-SEP-1999;	99US-0155339.
PR	PR	23-SEP-1999;	99US-0155486.
PR	PR	28-JUN-1999;	99US-0139763.
PR	PR	18-JUN-1999;	99US-0139459.
PR	PR	18-JUN-1999;	99US-0139451.
PR	PR	18-JUN-1999;	99US-0139456.
PR	PR	18-JUN-1999;	99US-0139462.
PR	PR	18-JUN-1999;	99US-0139463.
PR	PR	18-JUN-1999;	99US-0139458.
PR	PR	18-JUN-1999;	99US-0139492.
PR	PR	18-JUN-1999;	99US-0139450.
PR	PR	18-JUN-1999;	99US-0139454.
PR	PR	18-JUN-1999;	99US-0139455.
PR	PR	18-JUN-1999;	99US-0139119.
PR	PR	18-JUN-1999;	99US-0137528.
PR	PR	04-JUN-1999;	99US-0137502.
PR	PR	25-MAY-1999;	99US-0136629.
PR	PR	14-MAY-1999;	99US-0134370.
PR	PR	18-MAY-1999;	99US-0134668.
PR	PR	20-MAY-1999;	99US-0135124.
PR	PR	21-MAY-1999;	99US-0135353.
PR	PR	24-MAY-1999;	99US-0134219.
PR	PR	14-MAY-1999;	99US-0134221.
PR	PR	14-MAY-1999;	99US-0134287.
PR	PR	18-MAY-1999;	99US-0134941.
PR	PR	20-MAY-1999;	99US-0135124.
PR	PR	21-MAY-1999;	99US-0135353.
PR	PR	03-JUN-1999;	99US-0137528.
PR	PR	04-JUN-1999;	99US-0137502.
PR	PR	07-JUN-1999;	99US-0137374.
PR	PR	08-JUN-1999;	99US-0138094.
PR	PR	10-JUN-1999;	99US-0138540.
PR	PR	14-JUN-1999;	99US-0139277.
PR	PR	16-JUN-1999;	99US-0139499.
PR	PR	18-JUN-1999;	99US-0139499.
PR	PR	18-JUN-1999;	99US-0139499.
PR	PR	18-JUN-1999;	99US-0139453.
PR	PR	17-JUN-1999;	99US-0139492.
PR	PR	18-JUN-1999;	99US-0139454.
PR	PR	18-JUN-1999;	99US-0139455.
PR	PR	18-JUN-1999;	99US-0139456.
PR	PR	18-JUN-1999;	99US-0139457.
PR	PR	18-JUN-1999;	99US-0139458.
PR	PR	18-JUN-1999;	99US-0139459.
PR	PR	18-JUN-1999;	99US-0139460.
PR	PR	18-JUN-1999;	99US-0139461.
PR	PR	18-JUN-1999;	99US-0139462.
PR	PR	18-JUN-1999;	99US-0139463.
PR	PR	18-JUN-1999;	99US-0139470.
PR	PR	18-JUN-1999;	99US-0139470.
PR	PR	21-JUN-1999;	99US-0139817.
PR	PR	22-JUN-1999;	99US-0139899.
PR	PR	23-JUN-1999;	99US-0140353.
PR	PR	23-JUN-1999;	99US-0140354.
PR	PR	24-JUN-1999;	99US-0140695.
PR	PR	28-JUN-1999;	99US-0140695.
PR	PR	29-JUN-1999;	99US-0140991.
PR	PR	30-JUN-1999;	99US-0141287.
PR	PR	01-JUL-1999;	99US-0141842.
PR	PR	01-JUL-1999;	99US-0142154.
PR	PR	02-JUL-1999;	99US-0142055.
PR	PR	06-JUL-1999;	99US-0142390.

PR	16-JUL-1999; 9905-0144086.	PR	08-OCT-1999; 9905-015832.
PR	19-JUL-1999; 9905-0144325.	PR	12-OCT-1999; 9905-0158369.
PR	19-JUL-1999; 9905-0144332.	PR	13-OCT-1999; 9905-0159294.
PR	19-JUL-1999; 9905-0144333.	PR	13-OCT-1999; 9905-0159295.
PR	19-JUL-1999; 9905-0144334.	PR	14-OCT-1999; 9905-0159329.
PR	19-JUL-1999; 9905-0144335.	PR	14-OCT-1999; 9905-0159330.
PR	20-JUL-1999; 9905-0144352.	PR	14-OCT-1999; 9905-0159331.
PR	20-JUL-1999; 9905-0144632.	PR	14-OCT-1999; 9905-0159637.
PR	20-JUL-1999; 9905-0144884.	PR	14-OCT-1999; 9905-0159638.
PR	21-JUL-1999; 9905-0144814.	PR	18-OCT-1999; 9905-0159584.
PR	21-JUL-1999; 9905-0145086.	PR	21-OCT-1999; 9905-0160741.
PR	21-JUL-1999; 9905-0145088.	PR	21-OCT-1999; 9905-0160767.
PR	22-JUL-1999; 9905-0145085.	PR	21-OCT-1999; 9905-0160768.
PR	22-JUL-1999; 9905-0145087.	PR	25-OCT-1999; 9905-0161404.
PR	22-JUL-1999; 9905-0145089.	PR	25-OCT-1999; 9905-0161405.
PR	22-JUL-1999; 9905-0145151.	PR	25-OCT-1999; 9905-0161406.
PR	23-JUL-1999; 9905-0145145.	PR	25-OCT-1999; 9905-0161359.
PR	23-JUL-1999; 9905-0145218.	PR	26-OCT-1999; 9905-0161360.
PR	23-JUL-1999; 9905-0145222.	PR	26-OCT-1999; 9905-0161361.
PR	26-JUL-1999; 9905-0145276.	PR	28-OCT-1999; 9905-0161920.
PR	27-JUL-1999; 9905-0145913.	PR	28-OCT-1999; 9905-0161992.
PR	27-JUL-1999; 9905-0145911.	PR	28-OCT-1999; 9905-0161993.
PR	27-JUL-1999; 9905-0145919.	PR	29-OCT-1999; 9905-0162142.
PR	28-JUL-1999; 9905-0145951.		
PR	02-AUG-1999; 9905-0146336.		
PR	02-AUG-1999; 9905-0146388.		
PR	02-AUG-1999; 9905-0146389.		
PR	03-AUG-1999; 9905-0147038.		
PR	04-AUG-1999; 9905-0147204.		
PR	04-AUG-1999; 9905-0147302.		
PR	05-AUG-1999; 9905-0147260.		
PR	06-AUG-1999; 9905-0147303.		
PR	06-AUG-1999; 9905-0147416.		
PR	09-AUG-1999; 9905-0147493.		
PR	09-AUG-1999; 9905-0147935.		
PR	10-AUG-1999; 9905-0148171.		
PR	11-AUG-1999; 9905-0148319.		
PR	12-AUG-1999; 9905-0148341.		
PR	13-AUG-1999; 9905-0148565.		
PR	13-AUG-1999; 9905-0148664.		
PR	16-AUG-1999; 9905-0149368.		
PR	17-AUG-1999; 9905-0149175.		
PR	18-AUG-1999; 9905-014926.		
PR	20-AUG-1999; 9905-0149722.		
PR	20-AUG-1999; 9905-0149723.		
PR	23-AUG-1999; 9905-0149902.		
PR	23-AUG-1999; 9905-0149930.		
PR	25-AUG-1999; 9905-0150566.		
PR	26-AUG-1999; 9905-0150884.		
PR	27-AUG-1999; 9905-0151065.		
PR	27-AUG-1999; 9905-0151066.		
PR	27-AUG-1999; 9905-0151080.		
PR	31-AUG-1999; 9905-0151303.		
PR	01-SEP-1999; 9905-0151930.		
PR	07-SEP-1999; 9905-0152363.		
PR	10-SEP-1999; 9905-0153070.		
PR	13-SEP-1999; 9905-0153358.		
PR	15-SEP-1999; 9905-0154018.		
PR	16-SEP-1999; 9905-0154039.		
PR	22-SEP-1999; 9905-0154779.		
PR	23-SEP-1999; 9905-0155139.		
PR	24-SEP-1999; 9905-0155486.		
PR	24-SEP-1999; 9905-0155659.		
PR	29-SEP-1999; 9905-0156596.		
PR	04-OCT-1999; 9905-0157117.		
PR	05-OCT-1999; 9905-0157753.		
PR	06-OCT-1999; 9905-0157865.		
PR	07-OCT-1999; 9905-0158029.		

RESULT 12

Query ID	Match AAG57317	Score 69	Length 63	DB 21;	Local 69%	Similarity 85.7%	Pred. 0;	No. 3.4e+02;	Mismatches 1;	Matches 0;	Indels 0;	Gaps 0;
QY	1 NFRVTNF	7										
Db	115 NFRETNF	121										

Arabidopsis thaliana protein fragment SEQ ID NO: 73847.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 73847.

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; terminaison sequence.

KW OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-1999; 9905-0121825.

PR 05-MAR-1999; 9905-0123180.

PR 09-MAR-1999; 9905-0123548.

PR 23-MAR-1999; 9905-0125788.

PR 25-MAR-1999; 9905-0126264.

PR 29-MAR-1999; 9905-0126785.

PR 01-APR-1999; 9905-0127462.

PR 06-APR-1999; 9905-0128234.

PR 08-APR-1999; 9905-0128714.

PR 16-APR-1999; 9905-0128845.

PR 19-APR-1999; 9905-0130077.

PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-014352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0143632.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-014884.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-014814.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-014086.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-014192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-014145.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-014218.
PR	14-MAY-1999;	99US-0134219.	PR	23-JUL-1999;	99US-014224.
PR	14-MAY-1999;	99US-0134221.	PR	26-JUL-1999;	99US-0142276.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-014389.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0143918.
PR	19-MAY-1999;	99US-0134941.	PR	27-JUL-1999;	99US-0143919.
PR	20-MAY-1999;	99US-0135124.	PR	28-JUL-1999;	99US-0143951.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0143986.
PR	21-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0143388.
PR	25-MAY-1999;	99US-0136021.	PR	02-AUG-1999;	99US-0143389.
PR	27-MAY-1999;	99US-0136392.	PR	03-AUG-1999;	99US-0147038.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147204.
PR	01-JUN-1999;	99US-0137222.	PR	04-AUG-1999;	99US-0147302.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147192.
PR	04-JUN-1999;	99US-0137502.	PR	05-AUG-1999;	99US-0143319.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	99US-0138094.	PR	06-AUG-1999;	99US-0147416.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147305.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0143260.
PR	16-JUN-1999;	99US-0139457.	PR	12-AUG-1999;	99US-0143341.
PR	18-JUN-1999;	99US-0139458.	PR	13-AUG-1999;	99US-0143565.
PR	18-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0143929.
PR	18-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0146684.
PR	18-JUN-1999;	99US-0139454.	PR	09-AUG-1999;	99US-0143368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0143426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0143722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0143723.
PR	21-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0143929.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0143902.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-014930.
PR	18-JUN-1999;	99US-0139463.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139470.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139475.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151067.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0150566.
PR	01-JUL-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-015363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-015066.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-015070.
PR	30-JUN-1999;	99US-01411287.	PR	15-SEP-1999;	99US-0150758.
PR	12-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-015018.
PR	13-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-015039.
PR	14-JUL-1999;	99US-0142624.	PR	20-SEP-1999;	99US-0151779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-015139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-015486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-015659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-015659.
PR	13-JUL-1999;	99US-0143542.	PR	20-OCT-1999;	99US-015369.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-015117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0151753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-015865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-015029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-015232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-015369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-015293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-015294.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-015329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-015330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 29-OCT-1999; 99US-0162142.

Query Match 69.6%; Score 32; DB 21; Length 372;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NFRYTNF 7
 Db 160 NFRETNF 166

RESULT 13
 AAU30499 standard; Protein: 392 AA.
 AC AAU30499;
 XX DT 18-DEC-2001 (first entry)
 DE Novel human secreted protein #990.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 OS Homo sapiens.
 XX WO200179449-A2.
 XX PD 25-OCT-2001.
 XX PT 2002-269010/31.
 XX 16-APR-2001; 2001WO-US08656.
 XX PR 18-APR-2000; 2000US-0552929.
 XX PR 26-JAN-2001; 2001US-0770160.
 PA (HYSE-) HYSEO INC.
 XX PI Tang YF, Liu C, Dumanac RT;
 XX DR WPI; 2001-611725/70.

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU30499 standard; Protein: 392 AA.
 CC sequences of novel human secreted proteins of the invention.

SQ Sequence 392 AA;

Query Match 69.6%; Score 32; DB 22; Length 392;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FRYTNF 7
 Db 153 FKYTNF 158

RESULT 14
 ABB91033 standard; Protein: 402 AA.
 ID ABB91033
 XX AC ABB91033;
 XX DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 244.
 XX KW Herbicidal; plant; agriculture; herbicide.
 XX OS Arabidopsis thaliana.
 XX PN WO20010210-A2.
 XX PD 07-FEB-2002.
 XX PF 28-AUG-2001; 2001WO-EP09892.
 XX PR 28-AUG-2001; 2001WO-EP09892.
 XX PA (FARB) BAYER AG.
 XX PT tietjen K, Weidler M;
 XX DR WPI; 2002-269010/31.

PT identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant organisms -
 XX PS Claim 5; SEQ ID NO 244; 261pp + Sequence Listing; English.
 XX CC The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

SQ Sequence 402 AA;

Query Match 69.6%; Score 32; DB 23; Length 402;

Search completed: March 25, 2003, 08:19:42
 Job time : 27.1818 secs

Query	Match	Score	DB	Length	Start	End
PR	12-AUG-1999;	990US-0148341.				
PR	13-AUG-1999;	990US-0148565.				
PR	16-AUG-1999;	990US-0149684.				
PR	17-AUG-1999;	990US-0149368.				
PR	18-AUG-1999;	990US-0149175.				
PR	20-AUG-1999;	990US-0149426.				
PR	20-AUG-1999;	990US-0149722.				
PR	20-AUG-1999;	990US-0149723.				
PR	23-AUG-1999;	990US-0149929.				
PR	23-AUG-1999;	990US-0149902.				
PR	25-AUG-1999;	990US-0150566.				
PR	26-AUG-1999;	990US-0150884.				
PR	27-AUG-1999;	990US-0151065.				
PR	27-AUG-1999;	990US-0151066.				
PR	31-AUG-1999;	990US-0151303.				
PR	31-AUG-1999;	990US-0151438.				
PR	01-SEP-1999;	990US-0151930.				
PR	07-SEP-1999;	990US-0152363.				
PR	10-SEP-1999;	990US-0153070.				
PR	11-SEP-1999;	990US-0153758.				
PR	15-SEP-1999;	990US-0154018.				
PR	16-SEP-1999;	990US-0154039.				
PR	20-SEP-1999;	990US-0154779.				
PR	22-SEP-1999;	990US-0155139.				
PR	23-SEP-1999;	990US-0155486.				
PR	24-SEP-1999;	990US-0155659.				
PR	24-SEP-1999;	990US-0156458.				
PR	29-SEP-1999;	990US-0156596.				
PR	04-OCT-1999;	990US-0157117.				
PR	05-OCT-1999;	990US-0157753.				
PR	06-OCT-1999;	990US-0157865.				
PR	07-OCT-1999;	990US-0158029.				
PR	08-OCT-1999;	990US-0158232.				
PR	12-OCT-1999;	990US-0158369.				
PR	13-OCT-1999;	990US-0159293.				
PR	13-OCT-1999;	990US-0159294.				
PR	14-OCT-1999;	990US-0159325.				
PR	14-OCT-1999;	990US-0159330.				
PR	14-OCT-1999;	990US-0159331.				
PR	14-OCT-1999;	990US-0159637.				
PR	18-OCT-1999;	990US-0159638.				
PR	21-OCT-1999;	990US-0159584.				
PR	21-OCT-1999;	990US-0160741.				
PR	21-OCT-1999;	990US-0160767.				
PR	21-OCT-1999;	990US-0160768.				
PR	21-OCT-1999;	990US-0160770.				
PR	21-OCT-1999;	990US-0160814.				
PR	21-OCT-1999;	990US-016015.				
PR	22-OCT-1999;	990US-0160980.				
PR	22-OCT-1999;	990US-0160981.				
PR	22-OCT-1999;	990US-0160989.				
PR	25-OCT-1999;	990US-0161404.				
PR	25-OCT-1999;	990US-0161405.				
PR	25-OCT-1999;	990US-0161406.				
PR	26-OCT-1999;	990US-0161359.				
PR	26-OCT-1999;	990US-0161360.				
PR	26-OCT-1999;	990US-0161361.				
PR	28-OCT-1999;	990US-0161920.				
PR	28-OCT-1999;	990US-0161992.				
PR	28-OCT-1999;	990US-0161993.				
PR	29-OCT-1999;	990US-0162142.				

Query Match Score 69.6%; Score 32; DB 21; Length 511;
 Best Local Similarity 85.7%; Pred. No. 5, 3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NFRYTNF 7
 Db 76 NFRFTNF 82

GenCore version 5.1.4_ps-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw mode1

Run on: March 25, 2003, 08:22:27 ; Search time 8.45455 Seconds

(without alignments) 56.911 Million cell updates/sec

Title: US-09-646-532B-3

Perfect score: 46

Sequence: 1 NFRYTNFAX 9

Scoring table: BLOSUM2

Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA,*

```

1: /cgn2_6_ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6_ptodata/2/pubpaa/PCM_NEW_PUB.pep:*
3: /cgn2_6_ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6_ptodata/2/pubpaa/US05_PUBCOMB.pep:*
5: /cgn2_6_ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6_ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6_ptodata/2/pubpaa/PCMUS_PUBCOMB.pep:*
8: /cgn2_6_ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6_ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6_ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6_ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6_ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6_ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6_ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	32	69.6	40	12 US-10-001-843-127
2	32	69.6	927	12 US-10-003-356-8
3	32	69.6	1338	10 US-09-402-100-4
4	31	67.4	39	10 US-09-864-761-4546
5	31	67.4	187	10 US-09-811-284-228
6	31	67.4	311	10 US-09-920-177-15
7	31	67.4	359	10 US-09-928-175-10
8	31	67.4	383	10 US-09-928-175-8
9	31	67.4	646	10 US-09-928-175-13
10	31	67.4	682	10 US-09-928-175-12
11	31	67.4	694	10 US-09-920-175-8
12	31	67.4	713	9 US-09-965-536A-6
13	31	67.4	718	10 US-09-928-175-3
14	31	67.4	730	10 US-09-920-175-7
15	31	67.4	737	9 US-09-965-536A-2
16	31	67.4	754	10 US-09-920-175-2
17	31	67.4	1163	10 US-09-826-660-21
18	31	67.4	1193	10 US-09-873-873-30
19	65.2	456	10	US-09-925-300-1595

ALIGNMENTS

RESULT 1
US-10-001-843-127
; Sequence 127, Application US/10001843
; Patent No. US2003132255A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Lean
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and
FILE REFERENCE: DEX-0267
CURRENT APPLICATION NUMBER: US10/001,843
PRIORITY APPLICATION NUMBER: 60/249,982
PRIOR FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 218
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 127
LENGTH: 40
TYPE: PRT
; ORGANISM: Homo sapien
; US-10-001-843-127

Query Match	Score	DB	Length
Best Local Similarity	69 %;	12	40;
Matches	62.5 %;	Pred. No.	6.6;
5;	Conservative	Mismatches	1;
		Indels	0;
		Gaps	0;

OY 1 NFRYTNFA 8
Db 18 NIRFTNFS 25

RESULT 2
US-10-003-356-8
; Sequence 8, Application US/10003356
; Patent No. US200314618A1
; GENERAL INFORMATION:
; APPLICANT: Loh, Si
; APPLICANT: Holloway, James L.

; TITLE OF INVENTION: Human v2 vomeronasal Receptor
 ; FILE REFERENCE: 00-10/
 ; CURRENT APPLICATION NUMBER: US/10/003, 356
 ; CURRENT FILING DATE: 2001-11-15
 ; PRIORITY APPLICATION NUMBER: 60/252, 373
 ; PRIORITY FILING DATE: 2000-11-21
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 8
 ; LENGTH: 927
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: Chimeric receptor.
 ; US-10-003-356-8
 Query Match 69.6%; Score 32; DB 12; Length 927;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NFRYTNF 7
 Db 523 NFRSTNF 529

RESULT 3
 US-09-402-100-4
 ; Sequence 4, Application US/09402100
 ; Patent No. US2001019834A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Daewoong Pharmaceutical Co, LTD
 ; APPLICANT: Kim, Byung-O
 ; APPLICANT: Shin, Sung-Seup
 ; APPLICANT: Yu, Young-Hyo
 ; APPLICANT: Park, Myung-Hwan
 ; APPLICANT: Choi, Deok-Joon
 ; APPLICANT: Jung, Hyung-Jin
 ; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Heliotrope
 ; FILE REFERENCE: 0136/0G140
 ; CURRENT APPLICATION NUMBER: US/09/402, 100
 ; CURRENT FILING DATE: 1999-09-27
 ; EARLIER APPLICATION NUMBER: KR 97-11950
 ; EARLIER FILING DATE: 1997-03-31
 ; EARLIER APPLICATION NUMBER: KR 97-11951
 ; EARLIER FILING DATE: 1997-03-31
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO: 4
 ; LENGTH: 138
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: caga/CTXA2B Chimeric protein
 ; US-09-402-100-4

Query Match 69.6%; Score 32; DB 10; Length 1338;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NFRYTN 6
 Db 505 NFKYTN 510

RESULT 4
 US-09-864-761-45646
 Sequence 45646, Application US/09864761
 ; General Information:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aeonica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864, 761
 ; CURRENT FILING DATE: 2000-05-23
 ; PRIORITY APPLICATION NUMBER: US 60/180, 312
 ; PRIORITY FILING DATE: 2000-02-04
 ; PRIORITY APPLICATION NUMBER: US 60/207, 456
 ; PRIORITY FILING DATE: 2000-05-26
 ; PRIORITY APPLICATION NUMBER: US 09/632, 366
 ; PRIORITY FILING DATE: 2000-08-03
 ; PRIORITY APPLICATION NUMBER: GB 24263.6
 ; PRIORITY FILING DATE: 2000-10-04
 ; PRIORITY APPLICATION NUMBER: US 60/236, 359
 ; PRIORITY FILING DATE: 2000-09-27
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00664
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00667
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00669
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00671
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00672
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00661
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/00670
 ; PRIORITY FILING DATE: 2001-01-30
 ; PRIORITY APPLICATION NUMBER: US 60/234, 687
 ; PRIORITY FILING DATE: 2000-09-21
 ; PRIORITY APPLICATION NUMBER: US 09/608, 408
 ; PRIORITY FILING DATE: 2000-06-30
 ; PRIORITY APPLICATION NUMBER: US 09/774, 203
 ; PRIORITY FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO: 45646
 ; LENGTH: 39
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC011255.3
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.82
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.71
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.72
 ; OTHER INFORMATION: EXPRESSED IN TESTIS, SIGNAL = 0.72
 ; US-09-864-761-45646

Query Match 67.4%; Score 31; DB 10; Length 39;
 Best Local Similarity 71.4%; Pred. No. 10;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 FRNTNFA 8
 Db 7 VRYNNFA 13

RESULT 5
 US-09-811-824-228
 Sequence 228, Application US/09811284
 ; Patent No. US20020058306A1
 ; General Information:
 ; APPLICANT: Vogeli, Gabriel
 ; TITLE OF INVENTION: No. US20020058306A1 G Protein-Coupled Receptors
 ; FILE REFERENCE: 091670US1

CURRENT APPLICATION NUMBER: US/09/811,284
 CURRENT FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/189,783
 PRIOR FILING DATE: 2000-03-16
 PRIOR APPLICATION NUMBER: 60/189,907
 PRIOR FILING DATE: 2000-03-16
 PRIOR APPLICATION NUMBER: 60/189,918
 PRIOR FILING DATE: 2000-03-16
 PRIOR APPLICATION NUMBER: 60/189,960
 PRIOR FILING DATE: 2000-03-16
 PRIOR APPLICATION NUMBER: 60/189,917
 PRIOR FILING DATE: 2000-03-16
 PRIOR APPLICATION NUMBER: 60/192,945
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/192,916
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/192,923
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/192,933
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/192,830
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/192,234
 PRIOR FILING DATE: 2000-03-27
 PRIOR APPLICATION NUMBER: 60/192,155
 PRIOR FILING DATE: 2000-03-24
 PRIOR APPLICATION NUMBER: 60/192,935
 PRIOR FILING DATE: 2000-03-29
 NUMBER OF SEQ ID NOS: 258
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 228
 LENGTH: 187
 TYPE: PRY
 ORGANISM: Homo sapiens
 US-09-811-284-228

Query Match 67.4%; Score 31; DB 10; Length 187;
 Best Local Similarity 62.5%; Pred. No. 47;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NFRYTNAF 8
 |||| :|
 Db 278 NFRYCSVA 285

RESULT 7
 US-09-28-175-10
 ; Sequence 10, Application US/09928175
 ; Patent No. US20020123618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rogers, No. US20020123618A1
 ; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
 ; FILE REFERENCE: 00-1229
 ; CURRENT APPLICATION NUMBER: US/09/928,175
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/224,455
 ; PRIOR FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 5
 ; LENGTH: 383
 ; TYPE: PRY
 ; ORGANISM: Homo sapiens
 US-09-928-175-5

Query Match 67.4%; Score 31; DB 10; Length 383;
 Best Local Similarity 62.5%; Pred. No. 96;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NFRYTNAF 8
 |||| :|
 Db 350 NFRYCSVA 357

RESULT 9
 US-09-928-175-13

Query Match 67.4%; Score 31; DB 10; Length 311;
 Best Local Similarity 62.5%; Pred. No. 78;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 13 Application US/09928175
; Sequence 13, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Pasty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIORITY NUMBER: 60/224,455
; PRIORITY FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-928-175-13

Query Match 67.4%; Score 31; DB 10; Length 646;
Best Local Similarity 62.5%; Pred. No. 1.6e+02; Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NFRYVNFA 8
Db 278 NFRYCSYA 285

RESULT 10 Application US/09928175
; Sequence 12, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928,175
CURRENT FILING DATE: 2001-08-10
PRIORITY NUMBER: 60/224,455
PRIORITY FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 682
TYPE: PRT
ORGANISM: Homo sapiens
US-09-928-175-12

Query Match 67.4%; Score 31; DB 10; Length 682;
Best Local Similarity 62.5%; Pred. No. 1.7e+02; Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NFRYVNFA 8
Db 314 NFRYCSYA 321

RESULT 12 Application US/09965536A-6
; Sequence 6, Application US/09965536A-6
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINITER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
FILE REFERENCE: D0041MP
CURRENT APPLICATION NUMBER: US/09/965,536A
PRIORITY NUMBER: 60/235,713
PRIORITY FILING DATE: 2000-09-27
PRIORITY NUMBER: 60/261,781
PRIORITY FILING DATE: 2001-01-16
PRIORITY NUMBER: 60/306,605
PRIORITY FILING DATE: 2001-07-19
PRIORITY NUMBER: 60/310,436
PRIORITY FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 713
TYPE: PRT
ORGANISM: Homo sapiens
US-09-965-536A-6

RESULT 13 Application US/09928175-3
; Sequence 3, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and

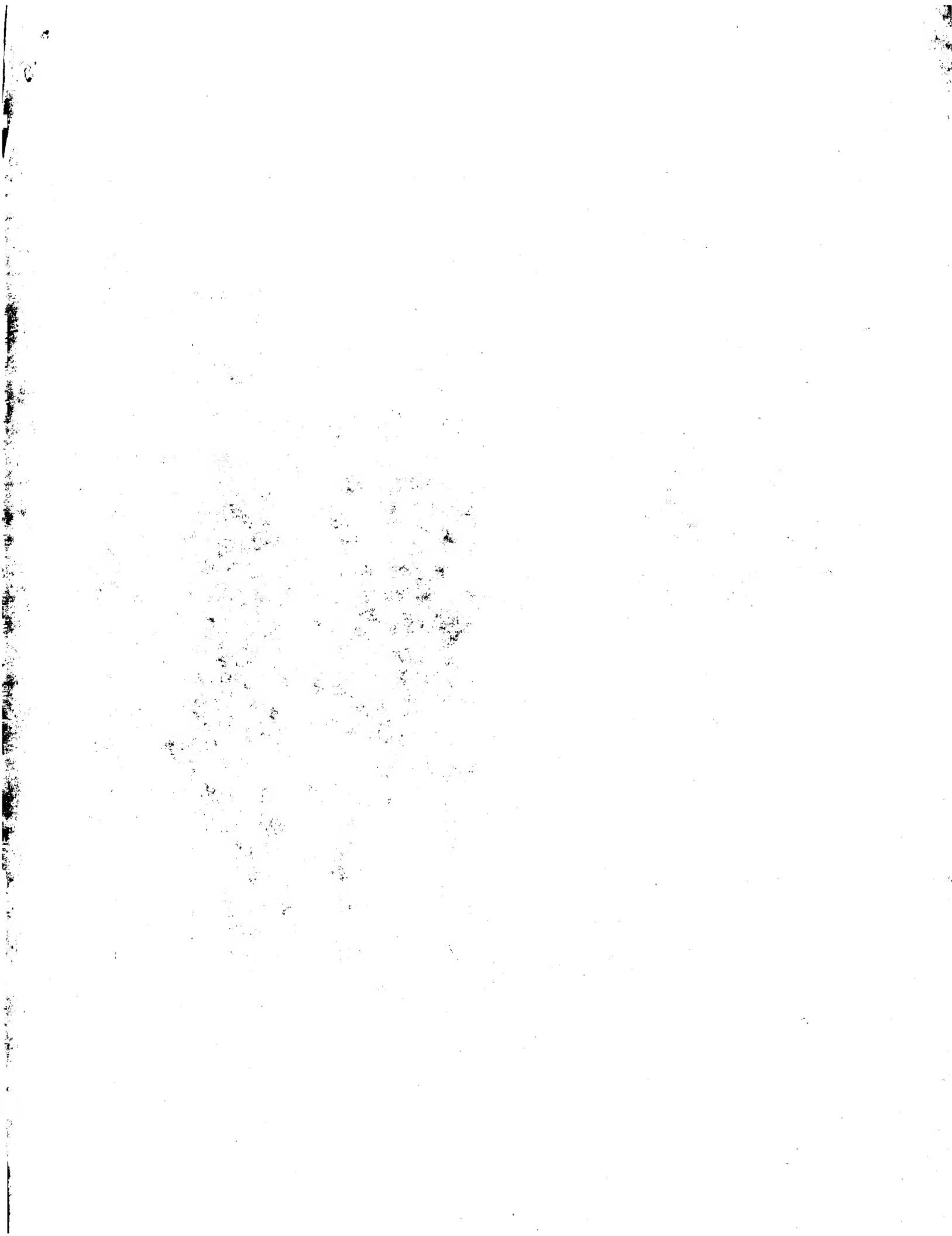
RESULT 14
 US-09-928-175-7
 ; Query Match 67.4%; Score 31; DB 10; Length 718;
 ; Best Local Similarity 62.5%; Pred. No. 1.8e+02; 2; Mismatches
 ; Matches 5; Conservative 2; Indels 0; Gaps 0;
 Qy 1 NFRYVNEA 8 |||| :||
 Db 350 NFRYCSYA 357

RESULT 14
 US-09-928-175-7
 ; Sequence 7, Application US/09928175
 ; Patent No. US20020123618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher J.
 ; APPLICANT: Daugherty, Betsy
 ; APPLICANT: Rogers, No. US2002123618Alma
 ; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
 ; FILE REFERENCE: 00-1229
 ; CURRENT APPLICATION NUMBER: US/09/928,175
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/224,455
 ; PRIOR FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 730
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-928-175-3

RESULT 14
 US-09-965-536A-2
 ; Query Match 67.4%; Score 31; DB 9; Length 737;
 ; Best Local Similarity 62.5%; Pred. No. 1.8e+02; 2; Mismatches
 ; Matches 5; Conservative 2; Indels 0; Gaps 0;
 Qy 1 NFRYVNEA 8 |||| :||
 Db 369 NFRYCSYA 376

Search completed: March 25, 2003, 08:23:42
 Job time: 10.4545 secs

RESULT 15
 US-09-965-536A-2
 ; Sequence 2, Application US/09965536A
 ; Publication No. US20030027333A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FEDER, J. N.
 ; APPLICANT: MINTIER, G.
 ; APPLICANT: RAMANATHAN, C. S.
 ; APPLICANT: HAWKEN, D. R.
 ; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
 ; FILE REFERENCE: DOO41NP
 ; CURRENT APPLICATION NUMBER: US/09/965, 536A
 ; CURRENT FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/235, 713
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/261, 781



Run on: March 25, 2003, 08:21:42 ; Search time 9.54545 Seconds
 (without alignments)
 27.742 Million cell updates/sec

Title: US-09-646-532B-3
 Perfect score: 46 NFRYINFA 9
 Sequence: BLOSUM62
 scoring table: Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents AA,*

1: /cgn2_6/podata/2/1aa/5A_COMB_pep:*

2: /cgn2_6/podata/2/1aa/5B_COMB_pep:*

3: /cgn2_6/podata/2/1aa/6A_COMB_pep:*

4: /cgn2_6/podata/2/1aa/6B_COMB_pep:*

5: /cgn2_6/podata/2/1aa/PCtUS_COMB_pep:*

6: /cgn2_6/podata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	38	82.6	28	6	Patent No. 5521285
2	32	69.6	245	4	Sequence 12, Appl
3	32	69.6	859	1	US-09-502-653-12
4	32	69.6	859	1	US-08-316-391B-2
5	32	69.6	859	2	US-09-034-06-2
6	32	69.6	859	4	US-09-259-437-2
7	32	69.6	859	5	PCT-US93-0982-2
8	32	69.6	1147	3	US-08-470-260-5
9	32	69.6	1147	3	US-08-471-491-5
10	32	69.6	1147	4	US-08-466-662-5
11	32	69.6	1181	1	US-08-316-397B-4
12	32	69.6	1181	2	US-09-034-06-4
13	32	69.6	1181	4	US-09-259-437-4
14	32	69.6	1181	5	PCT-US93-0982-4
15	32	69.6	1181	5	US-09-502-653-12
16	32	69.6	3289	2	US-08-477-451-2
17	31	67.4	9	1	US-08-188-277B-12
18	31	67.4	377	1	US-08-188-277B-2
19	31	67.4	377	1	US-08-188-277B-4
20	31	67.4	377	1	US-08-429-564-78
21	31	67.4	377	2	US-08-429-564-78
22	31	67.4	377	2	US-08-429-964-80
23	31	67.4	630	3	US-08-769-802A-2
24	31	67.4	630	3	US-08-461-551-2
25	31	67.4	823	1	US-08-461-551-2
26	31	67.4	823	4	US-09-37-621A-2
27	31	67.4	1030	4	US-09-091-117-2

ALIGNMENTS

RESULT 1
 5521285-3

; SEQ ID NO:3:
 ; LENGTH: 28

5521285-3

Query Match 82.6%; Score 38; DB 6; Length 28;
 Best Local Similarity 87.5%; Pred. No. 0.44;
 Matches 7; Conservative 0; Mismatches 1; Indels 0;
 QY |||||||
 1 NFRYINFA 8
 Db 15 NFDYTNFA 22

GENERAL INFORMATION:

APPLICANT: Bjarne Grønvald, Mads Eskelund
 APPLICANT: Clausen, Ib Groth
 APPLICANT: Schlein, Martin
 APPLICANT: Bech, Lisbeth
 APPLICANT: Stegaard, Peter Rahbek
 APPLICANT: Sjøholm, Carsten
 TITLE OF INVENTION: NOVEL GALACTANASES
 FILE REFERENCE: 5481-200-US
 CURRENT APPLICATION NUMBER: US-09-502-653
 CURRENT FILING DATE: 2000-02-11
 EARLIER APPLICATION NUMBER: PA 1999 00184
 EARLIER FILING DATE: 1999-02-11

EARLIER APPLICATION NUMBER: PA 1999 00799
 EARLIER FILING DATE: 1999-06-07
 EARLIER APPLICATION NUMBER: 601125, 885
 EARLIER FILING DATE: 1999-03-24
 EARLIER APPLICATION NUMBER: 601138, 445
 NUMBER OF SEQ ID NOS: 52
 SEQ ID NO: 12
 LENGTH: 245
 TYPE: PR
 ORGANISM: *Bacillus aquaraadhaerens* AC13 (DSM 8721)
 US-09-502-653-12

Query Match Similarity 69.6%; Score 32; DB 4; Length 245;
 Best Local Similarity 71.4%; Pred. No. 56; 1; Mismatches
 Matches 5; Conservative 1; Indels 0; Gaps 0;
 QY 1 NFRYRN 7
 |||:|||
 Db 47 NFHYSNF 53

RESULT 3

US-08-053-614-2

Sequence 2, Application US/08053614

Patent No. 5403924

GENERAL INFORMATION:

APPLICANT: COVER, TIMOTHY L.

APPLICANT: COVER, TIMOTHY L.

APPLICANT: BLASER, MARTIN J.

APPLICANT: BLASER, MARTIN J.

APPLICANT: TUMURU, MURALI K. R.

TITLE OF INVENTION: THE taga GENE AND METHODS FOR DETECTING

TITLE OF INVENTION: PREDISPOSITION TO PEPTIC ULCERATION

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

RESULT 4
 US-08-316-397B-2
 Sequence 2, Application US/08316397B
 ; Sequence 2, Application US/08316397B
 ; Patent No. 5733740
 ; GENERAL INFORMATION:
 ; APPLICANT: COVER, TIMOTHY L.
 ; APPLICANT: BLASER, MARTIN J.
 ; APPLICANT: HARRY KLEANTHOUS
 ; APPLICANT: TUMURU, MURALI K. R.
 ; TITLE OF INVENTION: THE taga GENE AND METHODS FOR DETECTING
 ; TITLE OF INVENTION: PREDISPOSITION TO PEPTIC ULCERATION AND GASTRIC CARCINOMA
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/316, 397B

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SPRATT, GWENDOLYN D.

REGISTRATION NUMBER: 36, 015

REFERENCE/DOCKET NUMBER: 2200.030

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404/688-0770

TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-316-397B-2

Query Match Similarity 69.6%; Score 32; DB 1; Length 859;
 Best Local Similarity 83.3%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0;
 QY 1 NFRYTN 6
 |||:|||
 Db 505 NFKYN 510

RESULT 5

US-09-034-3106-2

Sequence 2, Application US/090343106

Patent No. 5876943

GENERAL INFORMATION:

APPLICANT: COVER, TIMOTHY L.

APPLICANT: COVER, TIMOTHY L.

APPLICANT: BLASER, MARTIN J.

APPLICANT: BLASER, MARTIN J.

APPLICANT: HARRY KLEANTHOUS

APPLICANT: TUMURU, MURALI K. R.

TITLE OF INVENTION: THE taga GENE AND METHODS FOR DETECTING

TITLE OF INVENTION: PREDISPOSITION TO PEPTIC ULCERATION AND GASTRIC CARCINOMA

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

Query Match Similarity 69.6%; Score 32; DB 1; Length 859;
 Best Local Similarity 83.3%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0;
 QY 1 NFRYTN 6
 |||:|||
 Db 505 NFKYN 510

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/034,306
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/08/316,397
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: SPRATT, GWENDOLYN D.
 REGISTRATION NUMBER: 36,016
 TELECOMMUNICATION INFORMATION:
 REFERENCE/DOCKET NUMBER: 2200.030
 TELEPHONE: 404/688-0770
 FAX: 404/688-9880
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 859 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 IS-09-034-306-2
 Query Match 69.6%; Score 32; DB 2; Length 859;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NFRYTN 6
 Db 505 NFKYTN 510
 RESULT 7
 PCT-US93-09782-2
 Sequence 2, Application PC/TUS9309782
 GENERAL INFORMATION:
 APPLICANT: COVER, TIMOTHY L.
 APPLICANT: BLASER, MARTIN J.
 APPLICANT: HARRY KLEANTHous
 APPLICANT: TUMMURU, MURALI K.R.
 TITLE OF INVENTION: THE TAGA GENE AND METHODS FOR DETECTING
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NEEDLE & ROSENBERG, P.C.
 STREET: 133 Carnegie Way, Suite 400
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/09782
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: SPRATT, GWENDOLYN D.
 REGISTRATION NUMBER: 36,016
 REFERENCE/DOCKET NUMBER: 2200.009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404/688-0770
 FAX: 404/688-9880
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 859 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US93-09782-2
 Query Match 69.6%; Score 32; DB 5; Length 859;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NFRYTN 6
 Db 505 NFKYTN 510
 RESULT 8
 US-08-470-260-5
 Sequence 5, Application US/08470260
 Patent No. 607706
 GENERAL INFORMATION:
 APPLICANT: Covacci, Antonello
 ATTORNEY/AGENT INFORMATION:
 NAME: SPRATT, GWENDOLYN D.
 REGISTRATION NUMBER: 36,016
 REFERENCE/DOCKET NUMBER: 2200.030
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404/688-0770
 FAX: 404/688-9880

APPLICANT: Bugnoli, Massimo
 APPLICANT: Telford, John
 APPLICANT: Macchia, Giovanni
 TITLE OF INVENTION: Helicobacter Pylori Proteins Useful
 NUMBER OF SEQUNCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: California
 COUNTRY: USA
 ZIP: 94608-2916
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,260
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,848
 FILING DATE: 21-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0316.001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1147 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-470-260-5

Query Match 69.6%; Score 32; DB 3; Length 1147;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
 Db 505 NFRTYN 510

RESULT 9
 Sequence 5, Application US/08466662B
 ; PATENT NO. 6130059
 ; GENERAL INFORMATION:
 ; APPLICANT: Covacci, Antonello
 ; APPLICANT: Bugnoli, Massimo
 ; APPLICANT: Telford, John
 ; APPLICANT: Macchia, Giovanni
 ; APPLICANT: Rappoli, Rino
 ; TITLE OF INVENTION: Diagnostics
 ; FILE REFERENCE: CHIR057
 ; CURRENT APPLICATION NUMBER: US/08/466,662B
 ; CURRENT FILING DATE: 1995-06-06
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1147
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-08-466-662-5

Query Match 69.6%; Score 32; DB 4; Length 1147;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
 Db 505 NFRTYN 510

RESULT 11
 Sequence 4, Application US/08053614
 ; PATENT NO. 5403924
 ; GENERAL INFORMATION:
 ; APPLICANT: COVER, TIMOTHY L.
 ; APPLICANT: BLASER, MARTIN J.
 ; APPLICANT: TUMMURU, MURALI K.R.
 TITLE OF INVENTION: THE tagA GENE AND METHODS FOR DETECTING
 TITLE OF INVENTION: THE tagA GENE AND METHODS FOR DETECTING
 NUMBER OF SEQUNCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NEEDLE & ROSENBERG, P.C.
 STREET: 133 Carnegie Way, Suite 400
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/053,614
 FILING DATE: 19930426
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: SPRATT, GWENDOLIN D.
 REGISTRATION NUMBER: 36,016
 REFERENCE/DOCKET NUMBER: 2200.009

Query Match 69.6%; Score 32; DB 3; Length 1147;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
 Db 505 NFRTYN 510

RESULT 10
 Sequence 5, Application US/08466662B
 ; PATENT NO. 6130059
 ; GENERAL INFORMATION:
 ; APPLICANT: Covacci, Antonello
 ; APPLICANT: Bugnoli, Massimo
 ; APPLICANT: Telford, John
 ; APPLICANT: Macchia, Giovanni
 ; APPLICANT: Rappoli, Rino
 ; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/470,260
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: SPRATT, GWENDOLIN D.
 ; REGISTRATION NUMBER: 36,016
 ; REFERENCE/DOCKET NUMBER: 2200.009

Query Match 69.6%; Score 32; DB 3; Length 1147;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
 Db 505 NFRTYN 510

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404/688-0770

TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1181 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: Protein

US-08-053-614-4

Query Match

Best Local Similarity 69.6%; Score 32; DB 1; Length 1181;

Matches 5; Conservative 83.3%; Pred. No. 2.8e+02; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
Db 505 NFKYTN 510

RESULT 12

US-08-316-397B-4

Sequence 4, Application US/08316397B

PATENT NO. 573340

GENERAL INFORMATION:

APPLICANT: COVER, TIMOTHY L.

APPLICANT: BLASER, MARTIN J.

APPLICANT: HARRY KLEANTHOUS

APPLICANT: TUMMURU, MURALI K. R.

TITLE OF INVENTION: THE tagA GENE AND METHODS FOR DETECTING

TITLE OF INVENTION: PREDISPOSITION TO PEPTIC ULCERATION AND GASTRIC CARCINOMA

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/316,397

FILING DATE:

APPLICATION NUMBER: US/09/034,306

FILING DATE:

NAME: SPRATT, GWENDOLYN D.

CLASSIFICATION:

REFERENCE/DOCKET NUMBER: 36,016

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/316,397

FILING DATE:

NAME: SPRATT, GWENDOLYN D.

CLASSIFICATION:

REFERENCE/DOCKET NUMBER: 2200.030

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404/688-0770

TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1181 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-034-306-4

Query Match

Best Local Similarity 69.6%; Score 32; DB 2; Length 1181;

Matches 5; Conservative 83.3%; Pred. No. 2.8e+02; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
Db 505 NFKYTN 510

RESULT 14

US-09-259-437-4

Sequence 4, Application US/09259437

PATENT NO. 6153390

GENERAL INFORMATION:

APPLICANT: COVER, TIMOTHY L.

APPLICANT: BLASER, MARTIN J.

APPLICANT: HARRY KLEANTHOUS

APPLICANT: TUMMURU, MURALI K. R.

TITLE OF INVENTION: THE tagA GENE AND METHODS FOR DETECTING

TITLE OF INVENTION: PREDISPOSITION TO PEPTIC ULCERATION AND GASTRIC CARCINOMA

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

TELECOMMUNICATION INFORMATION:

SEQUENCE CHARACTERISTICS:

LENGTH: 1181 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-316-397B-4

Query Match

Best Local Similarity 69.6%; Score 32; DB 1; Length 1181;

Matches 5; Conservative 83.3%; Pred. No. 2.8e+02; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTN 6
Db 505 NFKYTN 510

RESULT 13

Copyright (c) 1993 - 2003 Compugen Ltd.	GenCore version 5.1.4-P5-4578				
On protein - protein search, using sw model					
Run on:	March 25, 2003, 08:20:22 ; Search time 9.81818 Seconds (without alignments)				
Perfect score:	46 88.123 Million cell updates/sec				
Sequence:	1 NFRYTNFAX 9				
Scoring table:	BLOSUM62				
Searched:	Gapop 10.0 , Gapext 0.5				
Total number of hits satisfying chosen parameters:	283224				
Minimum DB seq length:	0				
Maximum DB seq length:	200000000				
Post-processing:	Minimum Match 0% Listing first 45 summaries				
Database :	PIR_733:*				
	1: pir1:*				
	2: pir2:*				
	3: pir3:*				
	4: pir4:*				
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES					
Result NO.	Score	Query Match	Length	DB ID	Description
1	34	73.9	222	2 DB2132	hypothetical protein vc1981 [imported] - <i>Vibrio cholerae</i> (strain N16961 serogroup O1)
2	34	73.9	2802	2 F79686	cyclic beta (1-2) glucan
3	34	73.9	2831	2 A12911	beta (1->2) glucan
4	33	71.7	104	2 H97589	hypothetical protein
5	33	71.7	104	2 AG2811	hypothetical protein
6	33	71.7	157	2 T77883	major capsid protein
7	33	71.7	187	2 H81347	probable decarboxylase
8	33	71.7	262	2 T22371	hypothetical protein
9	33	71.7	437	1 VC1ECV	major capsid protein
10	33	71.7	511	2 F70110	conserved hypothetical protein
11	33	71.7	520	2 T18124	probable capsid protein
12	33	71.7	547	2 T47473	probable thiamin b
13	33	71.7	1315	2 T11055	membrane atpase -
14	32	69.6	115	2 S22588	V1 protein - tomato
15	32	69.6	258	2 T5361	hypothetical protein
16	32	69.6	315	2 A6944	outer membrane protein
17	32	69.6	346	2 C05950	hypothetical protein
18	32	69.6	369	2 E86307	similar to polygalacturonidase
19	32	69.6	467	2 T02238	glucosyl transferase
20	32	69.6	595	2 T05111	hypothetical protein
21	32	69.6	604	2 T08302	hypothetical protein
22	32	69.6	604	2 T02222	hypothetical protein
23	32	69.6	607	2 B75614	thiamin biosynthesis
24	32	69.6	713	2 T20090	hypothetical protein
25	32	69.6	1151	2 S48431	probable membrane protein
26	32	69.6	1167	2 BY1924	cag island protein
27	32	69.6	1186	2 C64588	cag pathogenicity
28	32	69.6	1215	2 B43281	cytotoxin-associated
29	32	69.6	1465	2 T23056	hypothetical protein
ALIGNMENTS					
Query Match	73.9%; Score 34; DB 2; Length 2802;				

A;Experimental source: serotype 02, strain NCIC 11168
 C;Genetics:
 A;Gene: CJ0768c
 C;Superfamily: defd protein

Query Match Best Local Similarity 71.7%; Score 33; DB 2; Length 187;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTNF 7
 ||::||:
 Db 56 NFKYTHF 62

RESULT 8

T22371 hypothetical protein F47H4.8 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T22371
 R;Steward, C.
 Submitted to the EMBL Data Library, March 1997
 A;Reference number: Z19555
 A;Accession: T22371
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-262 <WIL>
 A;Cross-references: EMBL:292786; PIDN:CAB07207.1; GSPDB:GN00023; CESP:F47H4 .8
 A;Experimental source: clone F47H4
 C;Genetics:
 A;Gene: CESP:F47H4.8
 A;Map position: 5
 A;Introns: 67/2; 98/2; 131/2; 186/1; 199/3

Query Match Best Local Similarity 71.7%; Score 33; DB 2; Length 187;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NFRYTNF 7
 ||::||:
 Db 206 NFKYANF 212

RESULT 9

VXECV major capsid protein - Chlorella virus PBCV-1
 C;Species: Chlorella virus PBCV-1
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 29-Oct-1999
 C;Accession: A40245; T17933
 R;Graves, M.V.; Meints, R.H.
 Virology 198, 198-207, 1992.
 A;Title: Characterization of the major capsid protein and cloning of its gene from algal
 A;Reference number: A40245; MUID:92230218; PMID:1566573
 A;Accession: A40245
 A;Molecule type: DNA
 A;Residues: 1-437 <GRA>
 A;Cross-references: GB:MB85052; NID:9323323; PIDN:AAA88828.1; PID:9323324
 R;Graves, M.V.; van Etten, J.L.
 Submitted to the EMBL Data Library, May 1999
 A;Reference number: Z18806

Query Match Best Local Similarity 71.7%; Score 33; DB 2; Length 187;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NFRYTNFA 8
 ||::||:
 Db 33 RYTNFA 38

RESULT 10

F70110 conserved hypothetical protein BB0086 - Lyme disease spirochete
 C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
 C;Accession: F70110
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wheson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu, N.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Nature 390, 580-586, 1997.
 A;Authors: Smith, H.O.; Venter, J.C.
 A;Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
 A;Reference number: A70100; MUID:9805943; PMID:9403685
 A;Accession: F70110
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-511 <KLE>
 A;Cross-references: GB:AE001121; GB:AE000783; NID:92687963; PIDN: AAC66471.1; PID:9268
 A;Experimental source: strain B31
 C;Superfamily: conserved hypothetical protein HII117

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 511;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NFRYTNFA 8
 ||::||:
 Db 423 NFRYENFA 430

RESULT 11

T18124 probable capsid protein A622L - Chlorella virus PBCV-1
 C;Species: Chlorella virus PBCV-1
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T18124
 R;Graves, M.V.; van Etten, J.L.
 Submitted to the EMBL Data Library, May 1999
 A;Reference number: T18806

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 511;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NFRYTNFA 8
 ||::||:
 Db 423 NFRYENFA 430

RESULT 12

T44743 probable thiamin biosynthesis protein thiC [imported] - *Mycobacterium leprae*
 C;Species: *Mycobacterium leprae*
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C;Accession: T44743
 R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

Query Match 71.7%; Score 33; DB 1; Length 437;

submitted to the EMBL Data Library, January 1998
 A;Reference number: Z22831
 A;Accession: T44743
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-547 <JAM>
 A;Cross-references: EMBL:AL035159; PIDN:CAA22712.1
 A;Experimental source: cosmid B145_0
 A;Gene: thiC
 C;Superfamily: thiamin biosynthesis protein thiC

Query Match	Score	DB	Length
Best Local Similarity	71.7%	2	547
Matches	75.0%	Pred. No.	62;
		Mismatches	1;
		Indels	0;
		Gaps	0;

QY 1 NFRWTNFA 8
 Db 281 SFLYTINFA 288

RESULT 13

T41055

membrane atpase - fission yeast. (*Schizosaccharomyces pombe*)

C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T41055
 R;Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A;Reference number: Z21967
 A;Accession: T41055
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1315 <HTL>
 A;Cross-references: EMBL:AL031324; PIDN:CAA20449.1; GSPDB:GN00068; SPDB:SPCC1672.11c
 A;Experimental source: strain 972n-; cosmid c1672
 C;Genetics:
 A;Map position: 3

Query Match Score 33; DB 2; Length 1315;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Query Match	Score	DB	Length
Best Local Similarity	71.4%	No.	1.5e+02
Matches	5	Pred.	
		Mismatches	1;
		Indels	0;
		Gaps	0;

QY 1 NFRYTINF 7
 Db 256 NFRVNY 262

RESULT 14

S22588

V1 protein - tomato yellow leaf curl virus

C;Species: tomato yellow leaf curl virus

C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
 C;Accession: S22588; S39233; S39209
 R;Rheyr, A.; Bendahmane, M.; Matzelt, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
 Nucleic Acids Res 19, 6763-6769, 1991
 A;Title: Tomato Yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite
 A;Reference number: S22588; MUID:92107660; PMID:1840676
 A;Accession: S22588
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-115 <KHE>
 A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43462.1; PID:g62212
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
 R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
 submitted to the EMBL Data Library, December 1993
 A;Reference number: S39233
 A;Accession: S39233
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-115 <CRE>

A;Cross-references: EMBL:Z28390; NID:g101671; PIDN:CAA82235.1; PID:g433850
 R;Moris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
 submitted to the EMBL Data Library, August 1993
 A;Description: High Similarity among the tomato yellow leaf curl virus isolates from
 A;Reference number: S39209
 A;Accession: S39209
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-115 <NOR>
 A;Cross-references: EMBL:Z225751; NID:g433655; PIDN:CAA81024.1; PID:g433656
 C;Superfamily: tomato yellow leaf curl virus V2 protein

Query Match	Score	DB	Length
Best Local Similarity	69.6%	No.	20;
Matches	85.7%	Pred.	
		Mismatches	0;
		Indels	1;
		Gaps	0;

QY 1 NFRYTNP 7
 Db 58 NRRYTNF 64

RESULT 15

T25361

hypothetical protein T27E7.5 - *caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T25361
 R;Cummings, P.
 submitted to the EMBL Data Library, November 1996
 A;Reference number: Z20023
 A;Accession: T25361
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-258 <WIL>
 A;Cross-references: EMBL:Z82284; PIDN:CAB05288.1; GSPDB:GN00022; CESP:T27E7.5
 A;Experimental source: clone T27E7
 C;Genetics:
 A;Gene: CESP-T27E7.5
 A;Map position: 4
 A;Introns: 5/3; 33/1; 142/3; 192/3; 211/1

Query Match Score 32; DB 2; Length 258;
 Best Local Similarity 69.6%; Pred. No. 46;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match	Score	DB	Length
Best Local Similarity	69.6%	No.	46;
Matches	5	Pred.	
		Mismatches	0;
		Indels	0;
		Gaps	0;

QY 1 NFRYTN 6
 Db 213 NYRTN 218

Search completed: March 25, 2003, 08:22:21
 Job time : 11.8182 secs

DT DT 01-JUL-1993 (Rel. 26; Last sequence update)
 DT DT 15-JUL-1998 (Rel. 36; Last annotation update)
 DE Major capsid protein (VP54).
 RA A430L
 RA Paramecium bursaria chlorella virus 1 (PBCV-1).
 RA Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OC NCBI_TaxID=10506;
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP MEDLINE=9223018; PubMed=1566573;
 RX Graves M.V.; Meints R.H.;
 RA Kutish G.F.; Li Y.; Lu Z.; Furuta M.; Rock D.L.; van Etten J.L.;
 RT "Characterization of the major capsid protein and cloning of its gene from algal virus PBCV-1;"
 RT Virology 188:198-207(1992).
 RL RN [2] SEQUENCE FROM N.A.
 RP MEDLINE=960190; PubMed=8806566;
 RX Graves M.V.; Meints R.H.;
 RA Kutish G.F.; Li Y.; Lu Z.; Furuta M.; Rock D.L.; van Etten J.L.;
 RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome; map positions 182 to 258;"
 RL Virology 223:303-317(1996).
 CC -1- SUBUNIT: HOMODIMER.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC DR EMBL: M85052; AAA88888.1; -
 DR EMBL: U42380; AAC96988.1; -
 DR PIR: A40245; VCEVCV.
 DR KW Coat protein; Glycoprotein; Late protein.
 FT INIT_MET 0
 FT DOMAIN 1 30 HYDROPHOBIC.
 FT CARBOND 46 46 N LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N LINKED (GlcNAc. . .) (POTENTIAL).
 FT CARBOHYD 369 369 N-LINKED (GlcNAc. . .) (POTENTIAL).
 FT CARBOHYD 375 375 N LINKED (GLCNAC. . .) (POTENTIAL).
 SQ 436 AA; 48034 MW; A71861A0BF8F813B CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RYTNFA 8
 DB 32 RYTNFA 37
 RESULT 3
 THIC_MYCLE STANDARD; PRT; 547 AA.
 ID Q9ZBG0;
 AC 09ZBG0;
 DT 30-MAY-2000 (Rel. 39; Created)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Thiamine biosynthesis protein thic.
 THIC OR MI_0294 OR MLCB1450.28C.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1769;
 SEQUENCE FROM N.A.
 RP STRAIN=TN;
 RC RN [1] SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V.; Gwilliam R.; Rajandream M.A.; Lyne M.; Lyne R.; Stewart A.;
 RA Sgouros J.; Peat N.; Hayles J.; Baker S.; Basham D.; Bowman S.;
 RA Brooks K.; Brown D.; Brown S.; Chillingworth T.; Churcher C.M.;
 RA Collins M.; Connor R.; Cronin A.; Davis P.; Feltwell T.; Fraser A.;
 RA Gentles S.; Goble A.; Hamlin N.; Harris D.; Hidalgo J.; Hodgson G.;
 RA Holroyd S.; Hornsby T.; Howarth S.; Huckle E.J.; Hunt S.; Jagesk K.;
 RA James K.; Jones L.; Jones M.; Leather S.; McDonald S.; McLean J.;
 RA Mooney P.; Moule S.; Mungall K.; Murphy L.; Niblett D.; Odell C.;
 RA Oliver K.; O'Neill S.; Pearson D.; Quail M.A.; Rabbinowitsch E.;
 RA Rutherford K.; Rutter S.; Saunders D.; Seeger K.; Sharp S.;
 RA Skelton J.; Simmonds M.; Squares R.; Squares S.; Stevens K.;
 RA Taylor K.; Taylor R.G.; Tivey A.; Walsh S.V.; Warren T.; Whitehead S.;
 RA Woodward J.; Volckaert G.; Kert R.; Robben J.; Gyromonoz B.;
 RA Weltjens I.; Vanstreels E.; Rieger M.; Schaefer M.; Mueller-Auer S.;
 RA Gabell C.; Fuchs M.; Fritzsche E.; Holzer E.; Moestl D.; Hilbert H.;
 RA

RA	Borzym K., Langer I., Beck A., Lehrauch H., Reinhardt R., Pohl T.M., RA	YVL-TYLCM	STANDARD;	PRT;	115 AA.
RA	Geffre P., Zimmermann W., Wedler H., Wambert R., Purnelle B., RA	ID-YVL-TYLCM	STANDARD;	PRT;	115 AA.
RA	Gallivert F., Aves S.J., Dreano S., Gloussy S., Lejaure M., Mottier S., RA	ID-P27270;			
RA	Dominguez R.R., Cruadado L., Jimenez J., Sanchez M., del Rey F., Benito J., RA	DT-01-AUG-1992 (Rel. 23, created)			
RA	Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., RA	DT-01-AUG-1992 (Rel. 23, last sequence update)			
RA	Schwarzki G.V., Ussey D., Barrell B.G., Nurse P., RT	DT-01-OCT-1994 (Rel. 30, last annotation update)			
RA	"The genome sequence of Schizosaccharomyces pombe.", Nature 413:871-880(2001);	DE-Hypothetical 13.3 kDa protein (VI protein).			
RL	-	VI.			
CC	- - CATALYTIC ACTIVITY: AMP + H(2)O = ADP + phosphate.	OS-Tomato yellow leaf curl virus (strain Marmande) (TYLCV), and			
CC	- - SUBCELLULAR LOCATION: Integral membrane protein.	OS-Tomato yellow leaf curl virus (strain Murcia) (TYLCV).			
CC	- - SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (EL-E2 ATPases). SUBFAMILY V.	OC-Viruses: ssDNA viruses; Geminiviridae; Begomovirus.			
CC	-	NCBI_TAXID=10833, 37139; [1]			
CC	-	RN-SEQUENCE FROM N.A.			
CC	-	RC-STRAIN=Mandane; MEDLINE=9210660; PubMed=1840676;			
DR	EMBL: AL031324; CAA20449.1; -.	RC-Khevry-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Creepi S.,			
DR	InterPro: IPR011757; ATPase_E1-E2.	RC-MEDLINE=9425636; Pubmed=8198442;			
DR	InterPro: IPR01454; Hlgnsase/Hydrolase.	RC-Noris E., Hidalgo E., Accotto G., Moriones E.,			
DR	Pfam: PF00122; E1-E2_ATPase; 1.	RC-Roncero B.; "Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus.";			
DR	Pfam: PF00702; Hydrolase; 1.	RL-Nucleic Acids Res. 19:6763-6769(1991). [2]			
DR	PRINTS; PRO0119; CATATPASE.	RN-SEQUENCE FROM N.A.			
DR	PROSITE; PS00134; ATPase_E1-E2; 1.	RP-STRAIN=Murcia; Arch. Virol. 135:165-170(1994).			
CC	KW-Hypothetical protein; Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.	CC-CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN 1 139.	CC-DOMAIN.			
FT	TRANSMEM 140 161.	CC-POTENTIAL.			
FT	DOMAIN 162 167.	CC-EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 168 190.	CC-POTENTIAL.			
FT	DOMAIN 191 324.	CC-CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 325 347.	CC-POTENTIAL.			
FT	DOMAIN 348 350.	CC-EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 351 369.	CC-POTENTIAL.			
FT	DOMAIN 370 531.	CC-CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 532 551.	CC-EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 552 564.	CC-POTENTIAL.			
FT	TRANSMEM 565 586.	CC-CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN 587 1072.	CC-POTENTIAL.			
FT	TRANSMEM 1073 1091.	CC-EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 1092 1099.	CC-POTENTIAL.			
FT	TRANSMEM 1100 1117.	CC-CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN 1118 1135.	CC-POTENTIAL.			
FT	TRANSMEM 1136 1159.	CC-EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 1160 1180.	CC-POTENTIAL.			
FT	TRANSMEM 1181 1203.	CC-CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN 1204 1216.	CC-EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 1217 1236.	CC-POTENTIAL.			
FT	DOMAIN 1237 1253.	CC-EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 1254 1274.	CC-POTENTIAL.			
FT	DOMAIN 1275 1315.	CC-CYTOPLASMIC (POTENTIAL).			
FT	MOD_RES 619 619.	CC-PHOSPHORYLATION (BY SIMILARITY).			
FT	METAL 1015 1015.	CC-MAGNESIUM (BY SIMILARITY).			
FT	METAL 1019 1019.	CC-METAL (BY SIMILARITY).			
SQ	SEQUENCE 1315 AA; 148786 MW; 081A550E59F2B0B5 CRC64;	RESULT 6			
Query Match	Best Local Similarity 71.7%; Score 33; DB 1; Length 115;	OMPPECOLI			
Matches 5;	Best Local Similarity 71.4%; Pred. No. 66; Mismatches 1; Indels 0; Gaps 0;	ID-OMPPECOLI	STANDARD;	PRT;	315 AA.
OY	1 NFRYTNF 7	AC-P34210;			
OY	1 NFRYTNF 7	DT-01-FEB-1994 (Rel. 28, Created)			
OY	1 NFRYTNF 7	DT-01-NOV-1995 (Rel. 32, Last sequence update)			
OY	1 NFRYTNF 7	DT-15-JUN-2002 (Rel. 41, Last annotation update)			
OY	1 NFRYTNF 7	DE-Outter membrane protease omppe precursor (EC 3.4.23.-).			
OY	1 NFRYTNF 7	GN-OMPPECPX			

DE Hypothetical protein KIAA0205.
 GN KIAA0205.
 RN Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 RC STRAIN-K12 / ATCC 12435;
 RX MEDILINE=94117370; PubMed=8288530;
 RA [1]
 RA kaufmann A.; Stierhof Y.-D.; Henning U.;
 RT "New outer membrane-associated protease of Escherichia coli K-12.";
 RL J. Bacteriol. 176:3549-367(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / CCR63;
 RA Shimizu H.; Saiochi Y.; Suda Y.; Uehara K.; Sampei G.; Mizobuchi K.;
 RT "Complete nucleotide sequence of the F plasmid: its implications for
 organization and diversification of plasmid genomes.";
 RT submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20029733; PubMed=10561486;
 RA Matsuo E.; Sampei G.; Mizobuchi K.; Ito K.;
 RT "The plasmid F CmpP protease, a homologue of OmpT, as a potential
 obstacle to E. coli-based protein production.";
 RL FEBS Lett. 461:6-8(1999).
 CC -I- FUNCTION: PROTEASE; ALSO ACTS AS A RECEPTOR FOR BACTERIOPHAGE OX2.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY R26.
 CC -I- ACYLTRANSFERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP001918; BRA97899.1; -.
 DR PIR; S37473; S37473.
 DR HSSP; P09169; 1178.
 DR MEROPS; A26_002; -.
 DR Ecogene; EG40042; ompP.
 DR InterPro; IPR000336; OmpTin.
 DR PRINTS; PRO0482; OmpTin.
 DR PRODOM; PD011585; OmpTin; 1.
 DR PROSITE; PS00834; OMPTIN_1; 1.
 DR PROSITE; PS00835; OMPTIN_2; 1.
 KW Hydrolase; Aspartyl protease; Outer membrane; Transmembrane; Signal;
 KW Phage recognition; Plasmid; Complete proteome.
 FT SIGNAL 1 23 OUTER MEMBRANE PROTEASE CMP.
 FT CHAIN 24 315 BY SIMILARITY.
 FT ACT_SITE 103 103 BY SIMILARITY.
 FT ACT_SITE 105 105 BY SIMILARITY.
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT ACT_SITE 232 232 BY SIMILARITY.
 SQ SEQUENCE 315 AA; 35499 MW; 45595973C6B2E149 CRC64;
 Query Match 69.6%; Score 32; DB 1; Length 370;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 FRYTNF 7
 Db 132 FKYTNF 137
 RESULT 8
 CGAL_HELPY ID CGAL_HELPY STANDARD; PRT; 1147 AA.
 AC P80200; DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytotoxicity associated immunodominant antigen (120 kDa protein).
 GN CAGA OR CAGL.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OC Helicobacter; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter; Proteobacteria; epsilon subdivision; Helicobacter group;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CCUG 17874 / NCTC 11638;
 RX MEDILINE=9329625; PubMed=8510329;
 RA Covacci A.; Consini S.; Bugnoli M.; Petracca R.; Burroni D.;
 RA Macchia G.; Massone A.; Papini E.; Xiang Z.; Figura N.;
 RA Rappoport R.;
 RT "Molecular characterization of the 128-kDa immunodominant antigen of
 Helicobacter pylori associated with cytotoxicity and quinolone
 ulcer.;"
 RT Proc. Natl. Acad. Sci. U.S.A. 90:5791-5795(1993).
 RL PRO; NCTC 11638;
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CCUG 17874 / NCTC 11638;
 RX MEDILINE=97121442; PubMed=8962108;

RA Cenini S., Lang C., Xiang Z., Crabtree J., Ghara P., RA
 RA Borodovsky M., Rappoport R., Covacci A.; RA
 RT "cag," a pathogenicity island of Helicobacter pylori, encodes type I- RA
 specific and disease-associated virulence factors."; RA
 proc. Natl. Acad. Sci. U.S.A. 93:14648-14653 (1996). RA
 RN SEQUENCE OF 181-190: 314-328; 366-377; 420-430; 494-506; 661-677; RA
 RP 900-914 AND 1662-1677. RA
 RA Herrmann V., Herrmann J., Kist M.; RA
 RL Submitted (APR-1993) to the SWISS-PROT data bank. RA
 CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT, RA
 CC OR FUNCTION OF THE CYTOTOXIN. RA
 CC -----RA
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration RA
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation - RA
 CC the European Bioinformatics Institute. There are no restrictions on its RA
 CC use by non-profit institutions as long as its content is in no way RA
 CC modified and this statement is not removed. Usage by and for commercial RA
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>) RA
 CC or send an email to license@isb-sib.ch). RA
 CC -----RA
 CC DR PRINTS; PRO1553; TYPE4SSCAGA. RA
 CC DR PIR: S48431; S48431. RA
 CC DR SGD; S0001310; NEOL. RA
 CC DR InterPro; IPR001757; ATPase_E1-E2. RA
 CC DR InterPro; IPR001454; Hydrolase/Hydrolase. RA
 CC DR Pfam; PF00122; E1-E2_Atpase; 1. RA
 CC DR Pfam; PF00702; Hydrolase; 1. RA
 CC DR PROSITE; PS00154; ATPASE_E1_E2; 1. RA
 CC KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding. RA
 CC FT DOMAIN 880 885 POLY-ASN. RA
 CC FT CONFLICT 320 320 G -> A (IN REF. 3). RA
 CC FT CONFLICT 325 325 P -> F (IN REF. 3). RA
 CC FT CONFLICT 328 328 R -> K (IN REF. 3). RA
 CC FT CONFLICT 426 426 K -> E (IN REF. 3). RA
 CC FT CONFLICT 429 429 Q -> E (IN REF. 3). RA
 CC FT CONFLICT 673 675 AQN -> TED (IN REF. 3). RA
 CC FT CONFLICT 901 901 A -> T (IN REF. 3). RA
 CC FT CONFLICT 903 903 Q -> E (IN REF. 3). RA
 CC FT CONFLICT 907 907 L -> P (IN REF. 3). RA
 CC FT CONFLICT 910 910 P -> R (IN REF. 3). RA
 CC FT CONFLICT 914 914 Q -> E (IN REF. 3). RA
 CC FT CONFLICT 1072 1072 P -> S (IN REF. 3). RA
 CC FT CONFLICT 1074 1074 S -> D (IN REF. 3). RA
 CC SQ SEQUENCE 1147 AA; 128013 MW: AB92770835F68490 CRC64; RA
 CC -----RA
 CC Query Match 69.6%; Score 32; DB 1; Length 1147; RA
 CC Best Local Similarity 83.3%; Pred. No. 91; RA
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0; RA
 CC -----RA
 CC Qy 1 NFRYTN 6 RA
 CC Db 505 NFKYTN 510 RA
 CC -----RA
 RESULT 9 RA
 ATC7_YEAST STANDARD; PRT: 1151 AA. RA
 ID ATC7_YEAST RA
 AC P40527; RA
 DT 01-FEB-1995 (Rel. 31, Created) RA
 DT 01-FEB-1995 (Rel. 31, Last sequence update) RA
 DT 15-JUN-2002 (Rel. 41, Last annotation update) RA
 DE Potential phospholipid-transferring ATPase 4 (EC 3.6.3.1). RA
 DE NEO1 OR YIL048W. RA
 OS Saccharomyces cerevisiae (Baker's yeast). RA
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; RA
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces. RA
 NCBI_TaxID=4932; RA
 RN [1] RA
 SEQUENCE FROM N/A. RA
 STRAN=S28C / AB972; RA
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D., RA
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A., RA
 Gentles S., Hamlin N., Horsnell T.S., Hunt S., Jagels K., Jones M., RA
 Louis E., Lye G., Moule S., Moule T., Odeil C., Pearson D., RA
 RA NCBI_TaxID=4932; RA
 RN [1] RA
 SEQUENCE FROM N/A. RA
 MEDLINE=99120557; PubMed=9923682; RA
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C., RA
 RA Smith D.R., Noonan B., Guild B.L., de Jonge B.L., Carmel G., RA
 RA Tummolo P.J., Caruso A., Urias Nickelsen M., Mills D.M., Ives C., RA
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F., RA
 RA Trister T.J.; RA

RT	"Genomic sequence comparison of two unrelated isolates of the human	Query	Match	69.6%	Score	32;	DB	1;	Length	1182;	
RT	gastric pathogen Helicobacter pylori.;	Best Local	Similarity	83.3%	Pred.	No.	94;				
RL	Nature 397:176-180(1999).	Matches	5;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,										
CC	OR FUNCTION OF THE CYTOTOXIN.										
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	Qy	1	NFRYTN 6							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	:									
CC	the European Bioinformatics Institute. There are no restrictions on its	Db	505	NFKYTN 510							
CC	use by non-profit institutions as long as its content is in no way										
CC	modified and this statement is not removed. Usage by and for commercial										
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).										
CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,	RESULT	12	CAGA_HELPY	STANDARD;	PRT;	1186	AA.			
CC	OR FUNCTION OF THE CYTOTOXIN.	ID	CAGA_HELPY	STANDARD;	PRT;	1186	AA.				
DR	EMBL; AE001483; AAD06073.1; -.	AC	P5580;								
DR	InterPro; IPR005165; Caga.	DT	01-Nov-1997	(Rel. 35, Created)							
DR	InterPro; IPR004355; IVSec_caga.	DT	01-Nov-1997	(Rel. 35, Last sequence update)							
DR	Pfam; PR03507; Caga; 1.	DT	16-Oct-2001	(Rel. 40, Last annotation update)							
KW	PRINTS; PRO1553; TYPE4SSCAGA.	DE	Cytotoxicity associated immunodominant antigen (120 kDa protein)	(CAG Pathogenicity island protein 26).							
FT	Antigen; Complete proteome.	DE	CAGA OR CAI OR CAG6 OR HP0547.								
DOMAIN	DOMAIN 246 249	GN	Helicobacter pylori (Campylobacter pylori).								
FT	DOMAIN 882 889	OS	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;								
SQ	SEQUENCE 1167 AA; 129729 MW; FDSE6B81CEBD0F2 CRC64;	OC	Helicobacter;								
		OX	NCBI-TaxID=210;								
		RN	{1}								
		RP	SEQUENCE FROM N.A.								
		RC	STRAIN=24695 / ATCC 700392;								
		RX	MEDLINE=97394467; PubMed=9252185;								
		RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,								
		RA	Fleischmann R.D., Ketchum K.A., Kleinj H.-P., Gill S., Dougherty B.A.,								
		RA	Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,								
		RA	Fleischmann R.D., Richardson D., Dodson R., Khalak H.G., Glodek A.,								
		RA	Loftus B., Richardson D., Ketchum K.A., Kleinj H.-P., Gill S., Dougherty B.A.,								
		RA	McKinney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,								
		RA	Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,								
		RA	Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,								
		RA	Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,								
		RA	Venter J.C.;								
		RT	"The complete genome sequence of the gastric pathogen Helicobacter								
		RT	pylori."								
		RL	Nature 388:539-547(1997).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
		CC	the European Bioinformatics Institute. There are no restrictions on its								
		CC	use by non-profit institutions as long as its content is in no way								
		CC	modified and this statement is not removed. Usage by and for commercial								
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).								
		CC	-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,								
		CC	OR FUNCTION OF THE CYTOTOXIN.								
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								

ID PG1_HUMAN STANDARD; PRT; 377 AA.

AC P33609;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 13-JUN-2002 (Rel. 41, Last annotation update)

DE Geranylgeranyl transferase type I beta subunit (EC 2.5.1.-) (Type I protein geranylgeranyltransferase beta subunit) (Ggerase-I-beta).

GN PGTIB.

OS Homo sapiens (Human).

OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Homo sapiens (Human).

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-brain;

RC MEDLINE=94148804; PubMed=8106351;

RX Zhang F.-L., Diehl R.E., Kohl N.E., Gibbs J.B., Girois B., Casey P.J., Omer C.A.; "cDNA cloning and expression of rat and human protein RAP1A AND RAP1B PROTEINS. -I- FUNCTION: CATALYZES THE TRANSFER OF A GERANYL-GERANYL MOIETY FROM GERANYLGERANYL PYROPHOSPHATE TO PROTEINS HAVING THE C-TERMINAL SEQUENCE CYS-ALIPHATIC-ALIPHATIC-X. ACTS ON THE RAC1, RAC2, RAP1A AND RAP1B PROTEINS.

CC -I- COFACTOR: Binds ONE ZINC ION (BY SIMILARITY).

CC -I- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT.

CC -I- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE BETA SUBUNIT FAMILY.

CC -I- SIMILARITY: CONTAINS 4 PFTB REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC DR EML; I24116; AAA17756_1; -

CC DR InterPro; IPR001330; Prenyltrans.

CC DR Pfam; PF00432; prenyltrans; 5.

CC KW Transferase; Prenyltransferase; Repeat; zinc.

FT REPEAT 144 186

FT REPEAT 193 234

FT REPEAT 245 284

FT REPEAT 284 333

FT METAL 269 269

FT METAL 271 271

FT METAL 321 321

SQ SEQUENCE 377 AA; 42396 MW; 565CD9B6C087A4DF CRC64;

Query Match 67.4%; Score 31; DB 1; Length 377;

Best Local Similarity 83.3%; Pred. No. 47;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FRYTNF 7

Db 286 FOYTNE 291

RESULT 14

Query Match 67.4%; Score 31; DB 1; Length 377;

Best Local Similarity 83.3%; Pred. No. 47;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FRYTNF 7

Db 286 FOYTNE 291

RESULT 15

CKAA_BACUF STANDARD; PRT; 753 AA.

AC CKAA_BACUF

DT 02321;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Pesticidal crystal protein cry20Aa (Insecticidal delta-endotoxin DE CryXX(A)) (Crystalline entomocidal protein) (86 kDa crystal protein).

GN CRY20AA OR CRYXX(A)

OS Bacillus thuringiensis (subsp. fukuokaensis).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=132265;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=9809459; PubMed=9406385;

RA Lee H.-K., Gill S.S.; "Molecular cloning and characterization of a novel mosquitoicidal protein gene from *Bacillus thuringiensis* subsp. *fukuokaensis*.";

RL APPL. ENVIRON. Microbiol. 63:4664-4670(1997).
 CC -!- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MOSQUITOS. ACTIVE AGAINST AEDES AEGYPTI AND
 CC CULEX QUINQUEFASCIATUS LARVAE.
 CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -!- PTM: HAS LOW MOSQUITOCIDAL ACTIVITY PROBABLY DUE TO RAPID
 CC PROTEOLYSIS TO INACTIVE 56-AND 43-KDA PROTEINS.
 CC -!- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -!- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U82218; AAB93476.1; -.
 DR InterPro; IPR001178; Endotoxin.
 DR Pfam; PF00555; endotoxin_1.
 DR Toxin; Sporulation.
 SQ SEQUENCE 753 AA; 86139 MW; CE3DED573EA41D66 CRC64;
 Query Match 67.4%; Score 31; DB 1; Length 753;
 Best Local Similarity 62.5%; Pred. No. 94;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 NFRYTNEFA 8
 |||:|||
 Db 591 NFKYATFA 598

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 25, 2003, 08:19:47 ; Search time 19.9091 Seconds

(without alignments)
93.145 Million cell updates/sec

Title:

US-09-646-532B-3

Perfect score:

46 NFRYTNFAK 9

Sequence:

BIOSUM62

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched:

671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:

671580

Minimum DB seq length:

0

Maximum DB seq length:

200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21;*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteria:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	36	78.3	453	10 Q9LT17
2	36	78.3	457	10 Q93T40
3	34	73.9	222	16 Q9K0L7
4	34	73.9	2831	16 Q8UBX0
5	34	73.9	2832	2 Q93N05
6	33	71.7	99	1 Q9UXD1
7	33	71.7	104	16 Q8UW48
8	33	71.7	157	12 Q98A34
9	33	71.7	262	5 Q9X129
10	33	71.7	432	12 Q90311
11	33	71.7	436	12 Q9WAU3
12	33	71.7	437	12 Q36189
13	33	71.7	437	12 Q9YJ22
14	33	71.7	475	12 Q8QNG3
15	33	71.7	511	16 Q51113
16	33	71.7	520	12 Q41104

17	33	71.7	521	12 Q90312	090312 chlorella v
18	33	71.7	950	5 Q963G0	0963G0 heliothis v
19	33	71.7	952	5 Q9NH88	09nh88 helicoverpa
20	33	71.7	1498	4 Q9P267	09p267 homo sapien
21	33	71.7	1960	5 Q9U0KB	09u018 plasmidium
22	33	71.7	2115	17 Q8TONO	Q8tono methanoscarc
23	32	69.6	83	12 Q8V588	Q8v588 tomato yell
24	32	69.6	83	12 Q8V584	Q8v584 tomato yell
25	32	69.6	83	12 Q8V582	Q8v582 tomato yell
26	32	69.6	83	12 Q8V580	Q8v580 tomato yell
27	32	69.6	83	12 Q8V578	Q8v578 tomato yell
28	32	69.6	83	12 Q8V575	Q8v575 tomato yell
29	32	69.6	115	12 Q8V612	Q8v612 tomato yell
30	32	69.6	209	12 Q8VAC5	Q8vac5 white spot
31	32	69.6	563	5 Q45856	Q45856 carnobabdi
32	32	69.6	270	11 Q8R1E1	Q8r1e1 mus musculus
33	32	69.6	346	16 Q98P02	Q98pu2 mycoplasma
34	32	69.6	369	10 Q9SH4	Q9shh4 arachidopsis
35	32	69.6	370	11 Q9IYX5	Q9iyx5 mus musculus
36	32	69.6	467	10 P93709	P93709 nicotiana t
37	32	69.6	563	17 Q8TP23	Q8tp23 methanoscari
38	32	69.6	587	5 Q8T2F9	Q8t2f9 dictyostelia
39	32	69.6	595	10 Q9SM02	Q9smg2 arabidopsis
40	32	69.6	599	10 Q9Abn1	Q9abn1 arabidopsis
41	32	69.6	604	17 P01989	P01989 halobacteri
42	32	69.6	604	17 P01989	P01989 halobacteri
43	32	69.6	607	16 Q9YK2	Q9yk2 deinococcus
44	32	69.6	713	5 P91982	P91982 cyanorhabdi
45	32	69.6	912	11 P07410	P07410 mus musculus

ALIGNMENTS

RESULT 1

Q9LT17 PRELIMINARY; PRT: 453 AA.
ID Q9LT17
AC Q9LT17;
DT 01-OCT-2000 ('TREMBLrel. 15', created)
DT 01-OCT-2000 ('TREMBLrel. 15', last sequence update)
DE Oxysterol-binding protein.

OS Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;
OC Spermatophytina; Magnoliophytina; eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI-TAXID=3702;

[1]

SEQUENCE FROM N.A.

STRAIN=COLUMBA;
RA Kaneko T., Kaoch T., Asamizu E., Sato S., Nakamura Y., Kotani H.,

Tabata S.;

"Structural analysis of Arabidopsis thaliana chromosome 5. XI.";

Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB025601; BAM974781; -;
DR InterPro; IPR000648; Oxysterol_BP.

DR Pfam; PF01237; Oxysterol_BP_1.

DR PROSITE; PS0013; OSBP; I.

SQ SEQUENCE: 453 AA; 51507 MW; FB326FFF00880D1EC CRC64;

Query Match Best Local Similarity Score 78.3%; DB 10; Length 453;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 NFRYTNFA 8
||:|||||
Db 323 NFQYTHFA 330

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

			RESULT 4
DT	01-DEC-2001 (TREMBLREL. 19, Created)	QSUBX0	PRELIMINARY;
DT	01-DEC-2001 (TREMBLREL. 19, Last sequence update)	QSUBX0	PRT; 2831 AA.
DE	Oxysterol-binding protein.	QSUBX0;	
GN	AT5G59350 OR F2015.10.	AC	
OS	Arabidopsis thaliana (Mouse-ear cress).	DT	01-JUN-2002 (TREMBLREL. 21, Created)
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	DT	01-JUN-2002 (TREMBLREL. 21, Last sequence update)
OC	Spermatophyta; Magnoliophyta; eudicots; Rosidae;	DT	01-MAR-2002 (TREMBLREL. 21, Last annotation update)
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.	DE	Beta (1-->2) glucan biosynthesis protein.
OX	NCBI_TAXID=3702;	GN	CHVB OR ATU2730 OR AGR_C1949.
RN	[1]	OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).
RP	SEQUENCE FROM N.A.	OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
RA	Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,	OC	Rhizobiaceae; Rhizobium.
RA	Palm C.J., Bowser L., Jones T., Banks J., Carninci P., Chen H.,	OX	NCBI_TAXID=176299;
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,	RN	[1]
RA	Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,	RP	SEQUENCE FROM N.A.
RA	Sakurai M., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,	RX	MEDLINE=21608550; PubMed=11743193;
RA	Ecker J., Theologis A., Davis R.W.,	RA	Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,
RA	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.	RA	Chen Y., Paulsen I.T., Eisen J.A., Harp P.D., Bovee D. Sr.,
DR	EMBL; AY054473; AAC96664.1; .	RA	Chapman P., Cleland J., Deatherage G., Gillett W., Grant C.,
DR	InterPro; IPR00048; oxysterol_BP_1.	RA	Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
DR	PF01237; Oxysterol_BP_1.	RA	Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
DR	PROSITE; PS01013; OSBP; UNKNOWN_1.	RA	Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespin W., Peery M.,
SEQUENCE	457 AA; 51966 MW; 90DDB0654110D85D CRC64;	RA	Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
QY	1 NFRYTNFA 8	RA	Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
QY	1 NFRYTNFA 8	RA	Nester E.W.,
Db	327 NFOYTHFA 334	RT	"the genome of the natural genetic engineer Agrobacterium tumefaciens
RT	C58.";	RL	Science 294:2317-2323(2001).
RN		RN	[2]
RP	SEQUENCE FROM N.A.	RX	MEMLINE=21608551; PubMed=11743194;
RX	Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,	RA	Qurolo B., Goldham B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA	Qurolo B., Goldham B.S., Cao Y., Akenazi M., Halling C., Mullin L.,	RA	Houmeli K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA	Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Marrelz B.,	RA	Flanagan C., Cowell C., Gurson J., Lomo C., Sear C., Scrub G.,
RA	Cielo C., Slater S.;	RA	"Genome sequence of the plant pathogen and biotechnology agent
RT	"Genome sequence of the plant pathogen and biotechnology agent	RT	Agrobacterium tumefaciens C58.";
RL	Science 294:2323-2328(2001).	RL	Science 294:2323-2328(2001).
DR	EMBL; AE009320; AAL4311.1; .	DR	EMBL; AE009320; AAL4311.1; .
DR	EMBL; AE008186; AAK8847.1; ALT_INTR.	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
SQ	SEQUENCE FROM N.A.	DR	Complete proteome.
RN	[1]	DR	SCIENCE 294:2317-2323(2001).
RP	SEQUENCE FROM N.A.	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RC	STRAINEL TOR NT6961 / SEROTYPE 01;	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RX	MEDLINE=20406833; PubMed=10952301;	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Podson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Gill S.R., Nelson K.E., Reed T.D., Tettelin H., Richardson D.A.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RA	Fraser C.M.,	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RT	cholerae.";	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
RL	Nature 406:477-483(2000).1; .	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
DR	TIGR; VC1981; .	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
KW	Hypothetical protein; Complete proteome.	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
SEQUENCE	222 AA; 24813 MW; 4FB63ED9E5214F2 CRC64;	DR	EMBL; AE008186; AAK8847.1; ALT_INTR.
QY	Query Match 73.9%; Score 34; DB 16; Length 222; Best Local Similarity 75.0%; Pred. No. 46; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	QY	1 NFRYTNFA 8
Db	84 NFRYTNFA 91	QY	1 NFRYTNFA 8
RA	Iannino F., Inon de Iannino N.;	RA	Iannino F., Inon de Iannino N.;

RT "Cyclic glucan synthetase of Agrobacterium tumefaciens, comparative sequence analysis.";
 RT Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 RL EMBL; AER391087; AAK7356.1;
 DR InterPro; IPR001999; Gram_Pos_anchor.
 DR InterPro; IPR001310; Zn_MtPeptrise.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING_UNKOWN_1.
 DR PROSITE; PS0142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 2832 AA; 316636 MW; BA525884BA3CLABA CRC64;

Query Match 73.9%; Score 34; DB 2; Length 2832;
 Best Local Similarity 71.4%; Pred. No. 5; $7e+02$; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY ||::|||
 DB 1397 NRYTYNF 7
 RN 1 NRQYTINF 1403

RESULT 6
 ID Q9UXD1 PRELIMINARY; PRT; 99 AA.
 AC Q9UXD1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Hypothetical 11.1 kDa protein.
 GN ORF-C20-035.
 SS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 RX NCBI_TaxID=2287;
 RN [1] SEQUENCE FROM N.A.

RC STRAIN=DSM 1617 / P2;
 RA Charlebois R.L., Singh R.K., Chan-Welther C.C.-Y., Allard G., Chow C.,
 RA Confortiieri F., Curtis B., Duquet M., Erauso G., Faguy D.,
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Koza C.,
 RA Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
 RA St. Jean A., Van Der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A.;
 RT "Gene content and organization of a 281-kbp contig from the genome of the extremely thermophilic archaeon, *Sulfolobus solfataricus* P2.";
 RT Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; Y18930; CAB57560.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 99 AA; 11090 MW; 83341F51627C1ABA CRC64;

Query Match 71.7%; Score 33; DB 1; Length 99;
 Best Local Similarity 71.4%; Pred. No. 33; Mismatches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NRQTYNF 7
 DB 47 NRFSNP 53

RESULT 7
 ID Q8UE48 PRELIMINARY; PRT; 104 AA.
 AC Q8UE48;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein Atu1913.
 GN Atu1913 OR AGR_C-3504.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 RN [1] NCBI_TaxID=176299;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;

RESULT 8
 ID Q98434 PRELIMINARY; PRT; 157 AA.
 AC Q98434;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE A382R protein.
 GN A382R.
 OS Paramecium buraria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 RX NCBI_TaxID=10506;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=96400190; PubMed=88065566;
 RX MEDLINE=2047854; PubMed=11021991;
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmer M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RA Lisee A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homospermidine
 RT synthase.";
 RL Virology 263:254-262(1999).
 RN [3] SEQUENCE FROM N.A.
 RP MEDLINE=2047854; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 PBCV-1.";
 RT Virology 276:27-36(2000).

RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

Query Match 71.7%; Score 33; DB 16; Length 104;
 Best Local Similarity 71.4%; Pred. No. 34; Mismatches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NRQTYHY 11
 DB 5 NRFSNP 11

RESULT 9
 ID Q98434 PRELIMINARY; PRT; 157 AA.
 AC Q98434;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE A382R protein.
 GN A382R.
 OS Paramecium buraria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 RX NCBI_TaxID=10506;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=96400190; PubMed=88065566;
 RX MEDLINE=2047854; PubMed=11021991;
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmer M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RA Lisee A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homospermidine
 RT synthase.";
 RL Virology 263:254-262(1999).
 RN [3] SEQUENCE FROM N.A.
 RP MEDLINE=2047854; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 PBCV-1.";
 RT Virology 276:27-36(2000).

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Kutayavim T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*";
 RT C58;"
 RL Science 294:2317-2323(2001).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Halling C., Mullin L.,
 RA Holum B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
 RA Holum K., Gordon J., Vaudin M., Lartchouk O., Epp A., Liu F.,
 RA William C., Allinger M., Dougherty D., Scott C., Lappas C., Mackellar B.,
 RA Flanagan C., Crowell C., Gurson J.J., Lomo C., Sear C., Strub G.,
 RA Cieci C., Slatier S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
 RL Science 294:2323-2326(2001).
 DR EMBL; AE009145; AAL2299.1; -.
 DR EMBL; AE008110; AAK87673.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 104 AA; 11642 MW; 8C5505A5BB69A308 CRC64;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karl P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deathrage G., Gillett W., Grant C.,
 RA Raymond C., Rouse G., Saephamachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestler E.W.;
 RT "The genome of

RN [4] SEQUENCE FROM N.A.
 RP Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
 RA van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
 RN [5] SEQUENCE FROM N.A.
 RP
 RA van Etten J.L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
 RN [7] SEQUENCE FROM N.A.
 RP
 RA Graves M.V.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
 RN [8] SEQUENCE FROM N.A.
 RP
 RA Graves M.V., van Etten J.L.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
 RN [9] SEQUENCE FROM N.A.
 RP
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: U045580; AAC98750.1; -
 SEQUENCE 157 AA; 18046 MW; 25EF8B6122E1014E CRC64;
 Query Match 71.7%; Score 33; DB 12; Length 157;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 RYTNEFA 8
 |||||
 Db 125 RYTNEFA 130

RESULT 9
 Q9XK29 PRELIMINARY; PRT; 262 AA.
 ID Q9XU29
 AC 09XU29;
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE F47H4.8 protein.
 GN F47H4.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderrinae; Caenorhabditida.
 OX NCBI_TaxID=6239;
 RN [1] SEQUENCE FROM N.A.
 RP
 RA Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
 RL
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=9906963; PubMed=9851916;
 RX MEDLINE=9902900; PubMed=10059963;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 investigating biology.";
 RL Science 262:2012-2018(1998).
 EMBL: Z92486; CAB07207.1; -.
 InterPro: IPR002900; DUF38.
 InterPro: IPR001810; F-box.
 DR Pfam: PF01827; DUF38; 1.
 DR SMART: SM00256; FB001; 1.
 DR PROSITE; PS50181; FB001; 1.
 SQ SEQUENCE 262 AA; 31280 MW; B8F55B2065558FC8 CRC64;

Query Match 71.7%; Score 33; DB 5; Length 262;
 Best Local Similarity 71.4%; Pred. No. 85;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NFRTNTNF 7
 ||:|| |||
 DB 205 NEKYANF 212

RESULT 10
 Q90311 PRELIMINARY; PRT; 432 AA.
 ID 090311;
 AC 090311;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Major capsid protein Vp49.
 GN Chlorella virus.
 OS Chlorella virus.
 OC viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 RX NCBI_TaxID=10307;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=CVG-1;
 RC MEDLINE=9922659; PubMed=10211977;
 RA Plugge B., Becker B., Wolf A.H.;
 RT "Several genes in Chlorella virus strain CVG-1 encode putative virion components.";
 RT J. Gen. Virol. 80:1067-1072(1999).
 DR EMBL: AF076921; AAC27492.1; -.
 DR InterPro: IPR002105; AATRNALIGASE_IT_1; UNKNOWN^1.
 DR PROSITE: PS00179; AA_TRNA_LIGASE_IT_1; UNKNOWN^1.
 SQ SEQUENCE 432 AA; 47696 MW; FD9732A43F876D1 CRC64;
 Query Match 71.7%; Score 33; DB 12; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYTNEFA 8
 |||||
 DB 33 RYTNEFA 38

RESULT 11
 Q9WAU3 PRELIMINARY; PRT; 436 AA.
 ID Q9WAU3
 AC 09WAU3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE Major capsid protein.
 GN CV2KCP2.
 OS Chlorella virus.
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OX NCBI_TaxID=10507;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=CVK2;
 RC MEDLINE=99170731; PubMed=10059963;
 RA Nishida K., Kinura Y., Kawashita T., Fujie M., Yamada T.;
 RT "Genetic variation of Chlorella viruses: Variable regions localized on the CVK2 genomic DNA.";
 RT Virology 255:76-84(1999).
 DR EMBL: AB018568; BAAT76600.1; -.
 SQ SEQUENCE 436 AA; 48398 MW; 01DAE208040C8D82 CRC64;

Query Match 71.7%; Score 33; DB 12; Length 436;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYTNEFA 8
 |||||
 DB 33 RYTNEFA 38

RESULT 12
 036189

ID	036189	PRELIMINARY;	PRT;	437 AA.
AC	036189;			
DT	01-JAN-1998 (TREMBLrel. 05; Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 08; Last annotation update)			
DE	Major capsid protein vp54.			
GN	vp54.			
OS	Chlorella virus.			
OS	Chlorella viruses, no RNA stage; Phycodnaviridae; Chlorovirus.			
OC	Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.			
OX	NCBI_TAXID=10507;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CIV2;			
RX	MEDLINE=98190351; PubMed=9514979;			
RA	Nishida K., Suzuki S., Kimura Y., Nomura N., Fujie M., Yamada T.;			
RT	"Group I introns found in Chlorella viruses: Biological implications."			
RT	Virology 242:319-326(1998).			
DR	EMBL: AB006978; BAA2198_1; -;			
DR	SEQUENCE 437 AA; 48049 MW; 53FFA9AE971F3683 CRC64;			
Query Match	71.7%; score 33; DB 12; Length 437;			
Best Local Similarity	100.0%; Pred. No. 1.4e+02;			
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	3 RYTNEA 8			
Db	33 RYTNEA 38			
RESULT	13			
09YZ8	PRELIMINARY;	PRT;	437 AA.	
ID	09YZ8;			
AC	09YZ8;			
DT	01-MAY-1999 (TREMBLrel. 10; Last sequence update)			
DT	01-MAY-1999 (TREMBLrel. 10; Last annotation update)			
DE	Major capsid protein vp52.			
GN	VP52.			
OS	Chlorella virus CVK2.			
OC	Chlorella virus K2, and			
OX	Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.			
RN	NCBI_TAXID=75912; 31555;			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=Chlorella virus K2;			
RA	Yamada T., Nishida K., Fujie M., Usami S.;			
RA	"Algal viruses use their own tRNAs expressed in host cells for the protein synthesis.";			
RT	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=Chlorella virus CVK2;			
RX	MEDLINE=9917071; PubMed=10069963;			
RA	Nishida K., Kimura Y., Kawasaki T., Fujie M., Yamada T.;			
RT	"Genetic variation of chlorella viruses: variable regions localized on the CVK2 genomic DNA."			
RL	Virology 255:386-384(1999).			
DR	EMBL: AB011506; BAA25143_1; -;			
DR	EMBL: AB018579; BAA76601_1; -;			
SO	SEQUENCE 437 AA; 48280 MW; ACAB83D9F235AF2D CRC64;			
Query Match	71.7%; score 33; DB 12; Length 437;			
Best Local Similarity	100.0%; Pred. No. 1.4e+02;			
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	3 RYTNEA 8			
Db	33 RYTNEA 38			
RESULT	14			
09YZ8	PRELIMINARY;	PRT;	511 AA.	
ID	09YZ8;			
AC	09YZ8;			
DT	01-JUN-1998 (TREMBLrel. 06; Created)			
DT	01-JUN-2002 (TREMBLrel. 21; Last sequence update)			
DE	Hypothetical protein BB0086.			
GN	BB0086.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TAXID=139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 35210 / B31;			
RX	MEDLINE=9806543; PubMed=9403685;			
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	Dougherty B.B., Tomb J.-F., Fleischmann R.D., Richardson D., Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J., Utterback T., Watney L., McDonald L., Artlach P., Bowman C., Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.,			
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."			
RT	Nature 390:580-586(1997).			
DR	EMBL: AE00121; AAC66471_1; -;			
DR	TIGR; BB0086; -;			
DR	InterPro; IPR004523; Mg_chelatse_chil.			
DR	InterPro; IPR004482; Mg_chelat_rel.			
DR	Pfam; PF01078; Mg_chelatase_1.			
DR	TIGRFAMS; TIGR0368; Mg_chelat_rel_1.			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 511 AA; 57237 MW; 18E7514CBCE6D0B5 CRC64;			
Query Match	71.7%; score 33; DB 16; Length 511;			
Best Local Similarity	75.0%; Pred. No. 1.7e+02;			
Matches	6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1 NFRYTNFA 8			
Db	423 NIRYENFA 430			

Fri Mar 28 10:51:43 2003

us-09-646-532b-3.rspt

Search completed: March 25, 2003, 08:21:37
Job time : 21.9091 secs

GenCore version 5.1.4-p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 08:17:33 ; Search time 40.7273 Seconds

(without alignments) 45.805 Million cell updates/sec

Title: US-09-646-532B-2

Perfect score: 62

Sequence: 1 GAXXXETAWACGXA 14

Scoring table: BL0S62

Gapop 10.0 , Gabext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq-101002:*

1: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1980.DAT:*

2: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1981.DAT:*

3: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1982.DAT:*

4: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1983.DAT:*

5: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1984.DAT:*

6: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1985.DAT:*

7: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1986.DAT:*

8: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1987.DAT:*

9: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1988.DAT:*

10: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1990.DAT:*

11: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1991.DAT:*

12: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1992.DAT:*

13: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1993.DAT:*

14: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1994.DAT:*

15: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1995.DAT:*

16: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1996.DAT:*

17: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1997.DAT:*

18: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1998.DAT:*

19: /SIDS1/gcadata/geneseq/geneseq-emb1/AAL1999.DAT:*

20: /SIDS1/gcadata/geneseq/geneseq-emb1/AK2000.DAT:*

21: /SIDS1/gcadata/geneseq/geneseq-emb1/AK2001.DAT:*

22: /SIDS1/gcadata/geneseq/geneseq-emb1/AK2002.DAT:*

23: /SIDS1/gcadata/geneseq/geneseq-emb1/AK2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1

AY4274 standard; peptide: 14 AA.

AY42754:

DT 20-DEC-1999 (first entry)

DE Wheat amyloplast ADP-glucose transporter peptide #2.

XX Starch biosynthesis; amyloplast; ADP-glucose; transport; import;

KW amlopectin; amylose; branching; chemical structure; transgenic plant;

KW optimisation; industrial applications.

XX Triticum aestivum.

OS location/Qualifiers

FH Key³ /label= Xaa

FT Misc-difference /note= "Xaa = any amino acid"

FT Misc-difference 4 /label= "Xaa = any amino acid"

FT Misc-difference 5 /note= "Xaa = any amino acid"

FT Misc-difference 13 /label= "Xaa = any amino acid"

FT Misc-difference 13 /note= "Xaa = any amino acid"

FT Misc-difference 13 /note= "Xaa = any amino acid"

FT Misc-difference 13 /note= "Xaa = any amino acid"

FT Misc-difference 13 /note= "Xaa = any amino acid"

FT Misc-difference 13 /note= "Xaa = any amino acid"

PD 23-SEP-1999.

11	50	80.6	25	23	AAE7106	Integrin-binding p
12	50	80.6	28	23	AAU74971	Alpha 5 beta 1 int
13	50	80.6	28	23	AAE7121	Integrin-binding o
14	50	80.6	31	23	AAU74976	Integrin binding o
15	44	71.0	10	23	AAU74982	Transfection assoc
16	44	71.0	10	23	AAE7112	Cyclic integrin-bi
17	44	71.0	12	23	AAU74961	Integrin-specific
18	44	71.0	12	23	AAU7100	Non-RGD containing
19	44	71.0	12	23	AAE7110	Cyclic integrin-bi
20	44	71.0	30	23	AAE7118	Integrin-binding o
21	40	64.5	769	22	ABG1454	Novel human diagno
22	39	62.9	669	22	AEE7035	Human transporter
23	38	61.3	9	16	AAU6196	Alpha5/Beta1 integr
24	38	61.3	9	19	AAW57200	RGD-containing pep
25	38	61.3	9	20	AAW56413	Chimeric adenovir
26	38	61.3	9	21	AAV56389	Peptide #5311 enco
27	38	61.3	9	23	AAE7986	Human ligand #6 at
28	37	59.7	38	22	ABG23080	Non-RGD containing
29	38	61.3	10	23	AAU74980	Transfection assoc
30	38	61.3	10	23	AAE7111	Cyclic integrin-bi
31	38	59.7	13	19	AAW56041	Peptide #5328 enco
32	37	59.7	38	22	ABG10705	Human Peptide enco
33	37	59.7	38	23	ABG6196	Human DTRP polype
34	37	59.7	433	22	ABG1629	Novel human diagno
35	37	59.7	38	22	ABM10909	Human bone marrow
36	37	59.7	38	22	AMR18729	Peptide #5163 enco
37	37	59.7	38	23	AMR1201	Peptide #5238 enco
38	37	59.7	38	23	ABG10705	Human Peptide enco
39	37	59.7	381	23	ABG6196	Human DTRP polype
40	41	58.1	76	23	ABQ4967	Novel human diagno
41	36	58.1	116	21	AGG2355	Human ORF protein
42	36	58.1	129	21	AGG2569	Human secreted pro
43	44	58.1	157	22	ABG08461	Zea mays protein f
44	36	58.1	157	22	ABG72946	Novel human diagno
45	36	58.1	157	22	ABG72946	Human olfactory re

XX	19-MAR-1999;	99WO-GB00728.
XX	20-MAR-1998;	98GB-0005939.
XX	(UYMA-) UNIV VICTORIA MANCHESTER.	
XX	Emes MJ, Tetlow IJ, Bowsher CG;	
XX	WPI; 1999-590977/50.	
PT	New transporter complex protein useful for modulating starch content in plants, especially useful in food production -	
XX	Claim 1; Page 3; 28pp; English.	
PS	This sequence represents a wheat amyloplast ADP-glucose transporter peptide, #2. The wheat amyloplast membrane and comprises at least two proteins; this sequence, along with peptides #1 (AY42753) and #3 (AY42755) are components of one of these proteins. Peptides #4-#7 (AY42756-Y42759) represent sequences within the second protein (AY42760). The sugar nucleotide ADP-glucose is the immediate substrate for starch synthesis, which occurs in the amyloplast; however, ADP-glucose is mainly synthesised outside the amyloplast in the cytoplasm. The ADP-glucose transporter is responsible for the import of ADP-glucose into the amyloplast and therefore plays a pivotal role in the regulation of starch synthesis. The transporter not only influences starch yield, but also quality as the starch synthases involved in amylose and amylopectin synthesis have different affinities for ADP-glucose. Variations in the chemical structure of starch are determined by the ratio of amylose to amylopectin, and by the degree of branching in amylopectin in the starch polymer. These variations can significantly alter the properties of starch. The ADP-glucose transporter complex is useful for generating transgenic plants in which the starch quality can be optimised for industrial applications in which starch is used. For example, transgenic plants which have an increased amylose content in starch are useful for production of starch with increased viscosity and gel strength, which prevents baked food going stale so quickly. Conversely, an increased amylopectin content in the starch produces waxy starch useful as thickening agents in food and coatings.	
XX	Sequence 14 AA;	
Query Match	87.1%; Score 54; DB 20; Length 14;	
Best Local Similarity	100.0%;	Pred. No. 0.0091;
Matches	14;	Conservative 0;
Oy	Indels 0;	Mismatches 0;
Db	Gaps 0;	
Qy	1 GAXXXETAWACGXA 14	
Db	1 GAXXXETAWACGXA 14	
RESULT 2		
ID	AY39332	
XX	AY39332 standard; peptide; 14 AA.	
AC	AY39332;	
XX	01-DEC-1999 (first entry)	
DE	ADP glucose transporter peptide 2.	
XX	ADP glucose transporter; transform plant cell; wheat; starch production; waxy starch; thickening agent; food; coating; increased viscosity; stale; gel strength; baked food.	
OS	Triticum aestivum.	
XX		
FH	Key difference Location/Qualifiers	
FT	Misc-difference 3 /note= "Any amino acid"	
FT	Misc-difference 4 /note= "Any amino acid"	
FT		
FT	Misc-difference 5 /note= "Any amino acid"	
FT	Misc-difference 13 /note= "Any amino acid"	
FT	/note= "Any amino acid"	
XX	W09947681-A1.	
PN		
XX	23-SEP-1999.	
XX	19-MAR-1999;	99WO-GB00727.
XX	20-MAR-1998;	98GB-0005939.
XX	(UYMA-) UNIV VICTORIA MANCHESTER.	
PA	Emes MJ, Tetlow IJ, Bowsher CG;	
PA	WPI; 1999-571841/48.	
PT	ADP glucose transporter protein used for modifying plant starch production	
XX	Claim 1; Page 15; 26pp; English.	
CC	Peptides AY39331-Y39337 are fragments of an ADP glucose transporter protein. The protein contains at least one of the peptide sequences, and is capable of ADP glucose transport. A DNA molecule encoding an ADP glucose transporter protein can be used to transform plant cells. The ADP glucose transporter protein can be used to regulate starch production from a plant. The plants can then be used to produce waxy starches that can be used as thickening agents in food and coatings. Alternatively the plants can be used to produce a starch with increased viscosity and gel strength, which can be incorporated in baked food which stays fresh for a longer length of time.	
CC	Sequence 14 AA;	
Query Match	87.1%; Score 54; DB 20; Length 14;	
Best Local Similarity	100.0%;	Pred. No. 0.0091;
Matches	14;	Conservative 0;
Oy	Indels 0;	Mismatches 0;
Db	Gaps 0;	
Qy	1 GAXXXETAWACGXA 14	
Db	1 GAXXXETAWACGXA 14	
RESULT 3		
ID	AAW95412	
XX	AAW95412 standard; peptide; 12 AA.	
AC	AAW95412;	
XX	18-MAR-1999 (first entry)	
DE	Integrin-specific peptide.	
XX	Integrin; transfection complex; integrin-binding; lipid; immunisation; antisense therapy; enzyme; therapeutic agent; immunogen; cystic fibrosis; cancer; viral infection; human immunodeficiency virus; cardiovascular; restenosis; leukaemia; asthma; glaucoma.	
OS	Synthetic.	
XX		
PF	29-MAY-1998; 98WO-GB01577.	
XX	W09854347-A1.	
PR	29-MAY-1997; 97GB-0011115.	
XX		
PA	(CHIL-) INST CHILD HEALTH.	

XX
PI Hart SL;
XX
DR WPI; 1999-045366/04.

XX
PT New integrin-targeting transfection complex including lipid - used to improve transfection efficiency for a very wide range of cells, useful in, e.g. antisense therapy and genetic immunisation
XX
PS Claim 11; Page 50; 70pp; English.

XX
CC The invention relates to an integrin-targeting transfection complex. The complex comprises a nucleic acid, an integrin-binding component, a polycationic nucleic acid-binding component and a lipid. The complexes are used for in vivo or in vitro transfection of cells, specifically: (i) for treatment or prevention of disease (in humans or other animals) caused by defective or deficient genes; (ii) for immunisation; (iii) for antisense therapy, and (iv) for protein production in host cells, e.g. of enzymes, therapeutic agents, vaccinating immunogens and diagnostic antigens. Typical of the diseases that can be treated or prevented are cystic fibrosis, cancer, viral infection (e.g. human immunodeficiency virus), cardiovascular disease (e.g. restenosis), leukaemia, asthma and glaucoma. Incorporation of the lipid into the complex increases transfection levels from 1-10 percent to over 50 percent. This effect is observed with all cell types tested including those that are resistant to transfection by most plasmid vectors. The complexes can carry large genes, up to 125 kb, e.g. an artificial chromosome. The present sequence represents a claimed example of an integrin-binding peptide used in the transfection complexes.
XX
SQ Sequence 12 AA:
Query Match 80.6%; Score 50; DB 20; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.037; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GAXXXETAWAGC 12
Db 1 GACRRETAWAGC 12

RESULT 4
ID AAE17101
ID AAE17101 standard; peptide; 12 AA.
AC AAE17101;
XX
DT 18-APR-2002 (first entry)
XX
DE Cyclic integrin-binding peptide 6.
XX
KW Integrin binding component; polycationic nucleic acid-binding component; lipid component; prophylaxis; immunisation; antisense therapy; asthma; cystic fibrosis; cancer; viral infection; human immunodeficiency virus; HIV infection; vaccine; neuroblastoma; bone marrow stem cell disorder; leukaemia; adjuvant immunotherapy; eye disease; glaucoma; restenosis; integrin-binding peptide; cyclic.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 3..11
FT /note= "Forms a cyclic structure"
PN WO200192542-A2.
XX
PD 06-DEC-2001.
XX
XX
PR 30-MAY-2001; 2001WO-GB02394.
XX
PR 30-MAY-2000; 2000GB-0013089.
PR 30-MAY-2000; 2000GB-0013089.
XX
01-MAY-2001; 2001US-287410P.

XX
PA (ICHI-) ICH PRODN LTD.
XX
PI Hart SL;
XX
DR WPI; 2002-139612/18.

XX
PT Complex for transfecting cell with nucleic acid for treating, preventing conditions caused by deficiency in a gene in humans, has nucleic acid, lipid, integrin binding and polycationic nucleic acid-binding components
XX
PS Claim 17; Page 6; 108pp; English.

CC The invention relates to integrin-targetting vectors having enhanced transfection activity. The vector complex comprises a nucleic acid, an integrin binding component, a polycationic nucleic acid-binding component and a lipid component. The integrin binding component comprise an integrin-binding element and a spacer element. Complex of the invention is useful for transfecting cells in vitro or in vivo with a nucleic acid, for treatment or prophylaxis of a condition caused in human or a non-human animal by a defect and/or a deficiency in a gene, immunisation and antisense therapy of a human or a non-human animal. It is useful for transfecting bronchial and lung epithelial and corneal endothelium for gene therapy for cystic fibrosis, asthma and various cancers and viral infections for example human immunodeficiency virus (HIV) infection. It is also useful as a vaccine or for therapy of neuroblastoma and the effective transfection of primary smooth muscle cells, cardiac myocytes and haematopoietic cells. Haematopoietic cell transfection enables gene therapy, gene vaccination and antisense therapy of diseases involving haematopoietic cells, including leukaemia and bone marrow stem cell disorders, for example transfection of a cytokine gene may be used for adjuvant immunotherapy. Transfection of corneal endothelium is useful for treatment of eye disease affecting the cornea or corneal organ transplants, for example in glaucoma. A gene that prevents the proliferation of cells in blood vessel walls is introduced using complex of the invention to reduce restenosis. The present sequence is cyclic integrin-binding peptide of the invention. This peptide is specific for alpha5beta1 integrin.
XX
SQ Sequence 12 AA:
Query Match 80.6%; Score 50; DB 23; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.037; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GAXXXETAWAGC 12
Db 1 GACRRETAWAGC 12

RESULT 5
ID AAR76190
ID AAR76190 standard; peptide; 13 AA.
AC AAR76190;
XX
DT 24-JAN-1996 (first entry)
XX
DE Integrin binding peptide #2.
XX
KW High affinity; integrin binding peptide; alphav/beta1; alphav/beta5; alphav/beta3; RGD; stable configuration; wound healing; osteoclast attachment; bone; angiogenesis; metastasis; tumour; smooth muscle cell migration.
XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT Disulfide-bond 3..11
XX
PN WO9514714-A1.
XX

XX	01-JUN-1995.	PR	22-NOV-1994; 94WO-US13542.
PF	XX	PR	22-FEB-1999; 99US-0119851.
XX	XX	PR	20-SEP-1999; 99US-0406781.
XX	XX	PA	(PROT-) PROTEINIX INC.
PR	04-AUG-1994; 94US-0286861.	XX	
PR	24-NOV-1993; 93US-0158001.	PI	Kenten JH, Roberts SF, Lebowitz MS;
XX	PA	XX	DR
PA	(LJOL-) LA JOLLA CANCER RES FOUND.	XX	WPI; 2000-565258/52.
PT	Koivunen E, Ruoslahti E;	XX	Novel compounds for modulating the ubiquitination of target proteins
XX	WPI; 1995-206899/27.	PT	comprising a ubiquitination recognition element-target protein element.
PT	PR	PT	fusion, useful for treating viral infections -
XX	PT	XX	Disclosure; Page 22; 10pp; English.
PS	PT	PS	The present invention describes a compound (1) for activating the
XX	CC	CC	ubiquitination (Ub)n of a target protein comprising a Ub n recognition
XX	CC	CC	(peptide) element (URE) covalently linked to a target protein (peptide)
XX	CC	CC	element (TPE). (1) can have antiviral, antiparasitic, hepatotropic,
XX	CC	CC	antiinflammatory and cytostatic activities. The compound of (1) may be
XX	CC	CC	used to treat a viral infection (especially hepatitis A, B, C or G, HIV-1
XX	CC	CC	or 2, Herpes CMV, rabies or Rouse sarcoma virus (RSV), parasitic
XX	CC	CC	infection, an infection caused by an eukaryotic organism in a mammal, to
XX	CC	CC	treat a tumour or to control pests. The compound may also be used to
XX	CC	CC	screen for target protein binding elements, to develop pesticides (e.g.
XX	CC	CC	insecticides, rodenticides, fungicides and herbicides) and to control
XX	CC	CC	gene expression (gene therapy). The present sequence represents an
XX	CC	CC	example of a target protein binding element given in the exemplification
XX	SQ	XX	of the present invention.
SQ	Sequence 13 AA;	Query Match 80.6%; Score 50; DB 16; Length 13;	Query Match 80.6%; Score 50; DB 21; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.04; Mismatches 0; Indels 0; Gaps 0;	Best Local Similarity 75.0%; Pred. No. 0.04; Mismatches 0; Indels 0; Gaps 0;	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAXXXETAWACG 12	Qy 1 GAXXXETAWACG 12	Qy 1 GAXXXETAWACG 12	Qy 1 GAXXXETAWACG 12
Db 1 GACRRETAWACG 12	Db 1 GACRRETAWACG 12	Db 1 GACRRETAWACG 12	Db 1 GACRRETAWACG 12
RESULT 6	RESULT 7	RESULT 7	RESULT 7
ID AAB20796	ID AAU74986	ID AAU74986	ID AAU74986
XX	XX	XX	XX
XX	AC	AC	AC
XX	AAB20796;	AAB20796;	AAB20796;
XX	DT 03-JAN-2001 (first entry)	DT 09-APR-2002 (first entry)	DT 09-APR-2002 (first entry)
XX	DE	DE	DE
XX	Transfection associated, integrin binding peptide #10.	Transfection associated, integrin binding peptide #10.	Transfection associated, integrin binding peptide #10.
XX	KW	KW	KW
XX	Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic;	Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic;	Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic;
XX	KW	KW	KW
XX	ophthalmological; vasoconstrictor; vaccine; gene therapy; transfection;	ophthalmological; vasoconstrictor; vaccine; gene therapy; transfection;	ophthalmological; vasoconstrictor; vaccine; gene therapy; transfection;
XX	KW	KW	KW
XX	cystic fibrosis; asthma; cancer; leukaemia; glaucoma; gene vaccination;	cystic fibrosis; asthma; cancer; leukaemia; glaucoma; gene vaccination;	cystic fibrosis; asthma; cancer; leukaemia; glaucoma; gene vaccination;
XX	KW	KW	KW
XX	anti-sense therapy; eye disease; corneal organ transplant; integrin; transfection; restenosis.	anti-sense therapy; eye disease; corneal organ transplant; integrin; transfection; restenosis.	anti-sense therapy; eye disease; corneal organ transplant; integrin; transfection; restenosis.
XX	OS	OS	OS
XX	Synthetic.	Synthetic.	Synthetic.
XX	FW	FW	FW
XX	Key Disulphide_bond 1..9	Location/Qualifiers	Key Disulphide_bond 1..9
XX	/note= "Cysteine residues linked by a disulfide bond to form a cyclic structure"		/note= "Cysteine residues linked by a disulfide bond to form a cyclic structure"
FT	Disulphide_bond 14..22		Disulphide_bond 14..22
FT			/note= "Cysteine residues linked by a disulfide bond to form a cyclic structure"
OS			
XX	W020047220-A1.	PN	W020047220-A1.
XX	XX	XX	XX
PD	17-AUG-2000.	PD	06-DEC-2001.
XX	XX	XX	XX
PF	11-FEB-2000; 2000WO-US03436.	PF	30-MAY-2001; 2001WO-GB02396.
XX	XX	XX	XX

PR 30-MAY-2000; 2000GB-0013089.
 PR 30-MAY-2000; 2000GB-0013090.
 PR 01-MAY-2001; 2001US-287410P.
 XX
 PA (ICHI-) ICH PRODN LTD.
 XX
 PR Hart SL;
 XX
 DR Hart SL;
 XX
 WPI; 2002-114355/15.

XX transfecing confluent cells with nucleic acid for gene therapy or gene vaccination, comprises contacting the cells with a receptor-targeted vector having the nucleic acid and an agent that disrupts cell-cell junctions -

PS Claim 17; Page 17; 11pp; English.

CC The invention describes transfecting (I) confluent cells or other slowly dividing or non-dividing cells that are in contact with each other, with a nucleic acid. The method comprises contacting the cells with a receptor-targeted vector comprising the nucleic acid and an agent that disrupts cell-cell junctions under conditions suitable to effect transfection. (II) is useful for transfecting bronchial and lung epithelium for gene therapy for cystic fibrosis, asthma and also various cancers and viral infections e.g. human immunodeficiency virus (HIV) infection. Haemopoietic cell transfection enables gene therapy, gene vaccination and anti-sense therapy of diseases involving haemopoietic cells, including leukaemia and bone marrow stem cell disorders. (III) transfection of corneal endothelium is useful for treatment of eye disease affecting the cornea or corneal organ transplants, for e.g. in glaucoma. A gene preventing cell proliferation in blood vessel walls is introduced using an integrin targeting transfection vector complex (III) to reduce restenosis. (II) is useful for intracellular transport and delivery of anti-sense oligonucleotides, which enables antiviral and cancer therapy and is effective in transporting large DNA molecules. This sequence represents a peptide that will permit cyclisation by disulfide bond formation. It lacks the conserved RGB amino acid integrin binding sequence, but binds integrins, allowing the nucleic acid to pass into the cell, described in the method of the invention.

SQ Sequence 23 AA:

Query Match	80.6%	Score	50;	DB	23;	Length	23;
Best Local Similarity	75.0%;	Pred.	No.	0.07;			
Matches	9;	Conservative					
Mismatches	0;						
Indels	3;						
Gaps	0;						

OY 1 GAXXXETAWAGC 12

Do 12 GACRRETAWAGC 23

RESULT 8

ABE1716 ID ABE1716 standard; peptide: 23 AA.
 XX AAE1716;
 XX AAE1716;

DE 18-APR-2002 (first entry)

XX Cyclic integrin-binding peptide #18.

XX Integrin binding component; polycationic nucleic acid-binding component; lipid component; prophylaxis; immunisation; antisense therapy; asthma; cystic fibrosis; cancer; viral infection; human immunodeficiency virus; HIV infection; vaccine; neviroblastoma; bone marrow stem cell disorder; leukaemia; adjuvant immunotherapy; eye disease; glaucoma; restenosis; KW integrin-binding peptide; cyclic.

OS Unidentified.

XX WO20012542-A2.

XX 06-DEC-2001.

XX Query Match

80.6%	Score	50;	DB	23;	Length	23;
Best Local Similarity	75.0%;	Pred.	No.	0.07;		
Matches	9;	Conservative				
Mismatches	0;					
Indels	3;					
Gaps	0;					

SQ Sequence 23 AA:

Query Match	80.6%	Score	50;	DB	23;	Length	23;
Best Local Similarity	75.0%;	Pred.	No.	0.07;			
Matches	9;	Conservative					
Mismatches	0;						
Indels	3;						
Gaps	0;						

OY 1 GAXXXETAWAGC 12

Do 12 GACRRETAWAGC 23

RESULT 9

AAW95415 ID AAW95415 standard; peptide: 25 AA.

XX AC AAW95416;
 XX DT 18-MAR-1999 (first entry)

XX DE Integrin-binding peptide.

XX Integrin; transfection complex; integrin-binding; lipid; immunisation; KW antisense therapy; enzyme; therapeutic agent; immunogen; cystic fibrosis; KW cancer; viral infection; human immunodeficiency virus; cardiovascular; KW restenosis; leukaemia; asthma; glaucoma; cyclic; circular.

XX OS Synthetic.

XX	FH	Key	Location/Qualifiers	FH	Key	Location/Qualifiers
XX	FH	Disulfide-bond	3..11 /note= "disulphide bridge"	FT	Disulfide-bond	3..11 /note= "Cysteine residues linked by a disulphide bond to form a cyclic structure"
XX	FT			FT		
XX	FT			FT		
PN	W09854347-A1.			FT		
XX	PD	03-DEC-1998.		FT		
XX	PA	29-MAY-1998;	98WO-GB01577.	XX	W0200192543-A2.	
XX	PR	29-MAY-1997;	97GB-0011115.	PD	06-DEC-2001.	
XX	PA	(CHIL-)	INST CHILD HEALTH.	XX	30-MAY-2001;	2001WO-GB02396.
XX	PI	Hart SL;		XX	30-MAY-2000;	2000GB-0013089.
XX	DR	WPI; 1999-045366/04.		PR	30-MAY-2000;	2000GB-0013090.
XX	PT	New integrin-targeting transfection complex including lipid - used to improve transfection efficiency for a very wide range of cells, useful in, e.g. antisense therapy and genetic immunisation		PR	01-MAY-2001;	2001US-287410P.
PS	Claim 12; Page 50;	70pp; English.		XX	PA	(ICH-)
XX	CC	The invention relates to an integrin-targeting transfection complex. The complex comprises a nucleic acid, an integrin-binding component, a polycationic nucleic acid-binding component and a lipid. The complexes are used for treatment or prevention of disease (in humans or other animals) (i) for treatment or prevention of disease (in humans or other animals) caused by defective or deficient genes; (ii) for illumination; (iii) for antisense therapy, and (iv) for protein production in host cells, e.g. of enzymes, therapeutic agents, vaccinating immunogens and diagnostic antigens. Typical of the diseases that can be treated or prevented are cystic fibrosis, cancer, viral infection (e.g. human immunodeficiency virus), cardiovascular disease (e.g. restenosis), leukaemia, asthma and glaucoma. Incorporation of the lipid into the complex increases transfection levels from 1-10 percent to over 50 percent. This effect is observed with all cell types tested including those that are resistant to transfection by most plasmid vectors. The complexes can carry large genes, up to 125 kb, e.g. an artificial chromosome. The present sequence represents a claimed example of an integrin-binding peptide used in the transfection complexes.		PA	ICH PRODN LTD.	
XX	CC	Sequence 25 AA;	SQ	PI	Hart SL;	
XX	CC	Query Match 80.6%; Score 50; DB 20; Length 25;	QY	XX	WPI; 2002-114355/15.	
XX	CC	Best Local Similarity 75.0%; Pred. No. 0.075; Mismatches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	Db	XX		
XX	CC	RESULT 10		PS	Disclosure; Page 15;	111pp; English.
XX	AC	AAU74968 standard; Peptide; 25 AA.		XX	Transferring confluent cells with nucleic acid for gene therapy or gene vaccination, comprises contacting the cells with a receptor-targeted vector having the nucleic acid and an agent that disrupts cell-cell junctions -	
XX	AC	AAU74968;		CC	The invention describes transfecting (I) confluent cells or other slowly dividing or non-dividing cells that are in contact with each other, with a nucleic acid. The method comprises contacting the cells with a receptor-targeted vector comprising the nucleic acid, and an agent that disrupts cell-cell junctions under conditions suitable to effect	
XX	DE	09-APR-2002 (first entry)		CC	transfection. (I) is useful for transfecting bronchial and lung epithelial for gene therapy for cystic fibrosis, asthma and also various cancers and viral infections e.g. human immunodeficiency virus (HIV) infection. Haemopoietic cell transfection enables gene therapy, gene vaccination and anti-sense therapy of diseases involving haemopoietic cells including leukaemia and bone marrow stem cell disorders	
XX	DE	Non-RGD containing alpha 5 beta 1 integrin binding peptide.		CC	transfection of corneal endothelium is useful for treatment of eye disease affecting the cornea or corneal organ transplants, for e.g. in glaucoma. A gene preventing cell proliferation in blood vessel walls is introduced using an integrin targeting transfection vector complex (II) to reduce restenosis. (II) is useful for intracellular transport and delivery of anti-sense oligonucleotides, which enables antiviral and cancer therapy and is effective in transporting large DNA molecules. This sequence represents a cyclic peptide which lacks the conserved RGD integrin binding amino acid sequence but can bind integrins to allow the nucleic acid to pass into the cell, described in the method of the invention.	
XX	OS	Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic; ophthalmological; vasotropics; vaccine; gene therapy; transfection; cystic fibrosis; asthma; leukaemia; glaucoma; gene vaccination; anti-sense therapy; eye disease; corneal organ transplant; integrin; transfection; restenosis; alpha 5 beta 1 integrin.		XX	Sequence 25 AA;	
XX	OS	Synthetic.	SQ	XX	RESULT 11	
XX	OS	Query Match 80.6%; Score 50; DB 23; Length 25;	QY	XX	Query Match 80.6%; Score 50; DB 23; Length 25;	
XX	OS	Best Local Similarity 75.0%; Pred. No. 0.075; Mismatches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	Db	XX	Best Local Similarity 75.0%; Pred. No. 0.075; Mismatches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
XX	OS	RESULT 11		AC	AAE17106	
XX	OS	AAE17106 standard; peptide; 25 AA.		AC	AAE17106;	
XX	OS	AAE17106		XX	18-APR-2002 (first entry)	

XX
 KW Integrin binding component; polycationic nucleic acid-binding component;
 KW lipid component; prophylaxis; immunisation; antisense therapy; asthma;
 KW cystic fibrosis; cancer; viral infection; human immunodeficiency virus;
 KW HIV infection; vaccine; neuroblastoma; bone marrow stem cell disorder;
 KW leukaemia; adjvant immunotherapy; eye disease; glaucoma; restenosis;
 KW integrin-binding peptide.
 OS Unidentified.
 XX
 PN WO200192542-A2.
 XX
 PD 06-DEC-2001.
 XX
 PR 30-MAY-2001; 2001WO-GB02394.
 XX
 PR 30-MAY-2000; 2000GB-0013089.
 PR 01-MAY-2001; 2001US-287410P.
 XX
 PA (ICHI-) ICH PRODN LTD.
 XX
 PR Hart SL;
 XX
 DR WPI; 2002-139612/18.
 XX
 PT Disclosure; Page 7; 108pp; English.
 XX
 PR Complex for transfecting cell with nucleic acid for treating,
 PT preventing conditions caused by deficiency in a gene in humans, has
 PR nucleic acid, lipid, integrin binding and polycationic nucleic
 PT acid-binding components.
 XX
 PS The invention relates to integrin-targetting vectors having enhanced
 CC transfection activity. The vector complex comprises a nucleic acid,
 CC an integrin binding component, a polycationic nucleic acid-binding
 CC component and a lipid component. The integrin binding component
 CC comprises an integrin-binding element and a spacer element. Complex
 CC of the invention is useful for transfecting cells in vitro or in
 CC vivo with a nucleic acid, for treatment or prophylaxis of a condition
 CC caused in human or a non-human animal by a defect and/or a deficiency
 CC in a gene, immunisation and antisense therapy of a human or a non-human
 CC animal. It is useful for transfecting bronchial and lung epithelium and
 CC corneal endothelium for gene therapy for cystic fibrosis, asthma and
 CC also various cancers and viral infections for example human
 CC immunodeficiency virus (HIV) infection. It is also useful as a vaccine
 CC or for therapy of neuroblastoma and the effective transfection of
 CC primary smooth muscle cells, cardiac myocytes and haematopoietic cells.
 CC Haematopoietic cell transfection enables gene therapy, gene vaccination
 CC and antisense therapy of diseases involving haematopoietic cells,
 CC including leukaemia and bone marrow stem cell disorders, for example
 CC transfaction of a cytokine gene may be used for adjvant immunotherapy.
 CC Transfection of corneal endothelium is useful for treatment of eye
 CC disease affecting the cornea or corneal organ transplants, for example
 CC vessel walls is introduced using complex of the invention to reduce
 CC restenosis. The present sequence is integrin-binding peptide
 CC of the invention. This peptide is specific for alpha5beta1 integrin.
 XX
 SQ Sequence 25 AA:
 Query Match 80.6%; Score 50; DB 23; Length 25;
 Best Local Similarity 75.0%; Pred. No. 0.075;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY |||||
 OY 1 GAXXETAWAGC 12
 1 GACRTRPAWAGC 12

RESULT 12
AAU74971

RESULT 12
AAU74971

Query Match 80 6%; Score 50; DB 23; Length 28;
 Best Local Similarity 75 0%; Pred. No. 0.084; Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GAXXETAWAGC 12
 ||| |||||
 Db 17 GACRRETAWAGC 28

RESULT 13

AAE17121
 ID AAE17121 standard; peptide: 28 AA.
 XX
 AC
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Integrin-binding oligolysine-peptide 6.

XX
 KW Integrin binding component; polycationic nucleic acid-binding component;
 KW lipid component; prophylaxis; immunisation; antisense therapy; asthma;
 KW cystic fibrosis; cancer; viral infection; human immunodeficiency virus;
 HIV infection; vaccine; neuroblastoma; bone marrow stem cell disorder;
 KW leukaemia; adjuvant; immunotherapy; eye disease; glaucoma; restenosis;
 KW integrin-binding peptide.

XX
 OS unidentified.

XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 19..27 /note= "Forms a cyclic structure"
 XX
 PN WO200192542-A2.
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-GB02394.
 XX
 PR 30-MAY-2000; 2000GB-0013089.
 PR 01-MAY-2001; 2001US-287410P.
 XX
 PA (ICHI-) ICH PRODN LTD.
 XX
 PI Hart SL;
 XX
 DR WPI; 2002-139612/18.

XX
 PT Complex for transfecting cell with nucleic acid for treating,
 PT preventing conditions caused by deficiency in a gene in humans, has
 PT nucleic acid, lipid, integrin binding and polycationic nucleic
 acid-binding components -

XX
 PS Example 13; Page 50; 108pp; English.

XX
 CC The invention relates to integrin-targetting vectors having enhanced
 transfection activity. The vector complex comprises a nucleic acid,
 an integrin binding component, a polycationic nucleic acid-binding
 component and a lipid component. The integrin binding component
 comprises an integrin-binding element and a spacer element. Complex
 of the invention is useful for transfecting cells in vitro or in
 vivo with a nucleic acid, for treatment or prophylaxis of a condition
 caused in human or a non-human animal by a defect and/or a deficiency
 in a gene, immunisation and antisense therapy of a human or a non-human
 animal. It is useful for transfecting bronchial and lung epithelium and
 corneal endothelium for gene therapy for cystic fibrosis, asthma and
 also various cancers and viral infections for example human
 immunodeficiency virus (HIV) infection. It is also useful as a vaccine
 or for therapy of neuroblastoma and the effective transfection of
 primary smooth muscle cells, cardiac myocytes and haematopoietic cells.
 CC Haematopoietic cell transfection enables gene therapy, gene vaccination
 and antisense therapy of diseases involving haematopoietic cells,

CC including leukaemia and bone marrow stem cell disorders', for example
 CC transfection of a cytokine gene may be used for adjuvant immunotherapy.
 CC Transfection of corneal endothelium is useful for treatment of eye
 CC disease affecting the cornea or corneal organ transplants, for example
 CC in glaucoma. A gene that prevents proliferation of cells in blood
 CC vessel walls is introduced using complex of the invention to reduce
 peptide of the invention.

CC Sequence 28 AA;

RESULT 14

AAU74976
 ID AAU74976 standard; Peptide: 31 AA.
 XX
 AC AAU74976;
 XX
 DT 09-APR-2002 (first entry)

XX
 DE Integrin binding oligo lysine peptide #2.

XX
 KW Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic;
 KW ophthalmological; vasotonic; vaccine; gene therapy; transfection;
 KW cystic fibrosis; asthma; cancer; leukaemia; glaucoma; gene vaccination;
 KW anti-sense therapy; eye disease; corneal organ transplant; integrin;
 XX
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Region 1..16 /note= "Polycationic nucleic acid binding sequence"
 FT Region 17..21
 FT Misc-difference 17 /label= Spacer
 FT Misc-difference 19 /label= Unknown
 FT Disulfide-bond 22..30 /note= "Cysteine residues linked by a disulphide
 FT bond to form a cyclic structure"
 XX
 PN WO200192543-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-GB02396.
 XX
 PR 30-MAY-2000; 2000GB-0013089.
 PR 01-MAY-2001; 2001US-287410P.
 XX
 PA (ICHI-) ICH PRODN LTD.
 XX
 PI Hart SL;
 XX
 DR WPI; 2002-114355/15.

XX
 PT Transferring confluent cells with nucleic acid for gene therapy or gene
 PT vaccination, comprises contacting the cells with a receptor-targeted
 PT vector having the nucleic acid and an agent that disrupts cell-cell
 PT junctions -

XX
 PS Example 13; Page 55; 111pp; English.

XX
 CC The invention describes transfecting (I) confluent cells or other slowly
 CC dividing or non-dividing cells that are in contact with each other, with
 CC a nucleic acid. The method comprises contacting the cells with a
 CC receptor-targeted vector comprising the nucleic acid, and an agent that
 CC disrupts cell-cell junctions under conditions suitable to effect
 CC transfection. (II) is useful for transfecting bronchial and lung
 CC epithelium for gene therapy for cystic fibrosis, asthma and also various
 CC cancers and viral infections e.g. human immunodeficiency virus (HIV)
 CC infection. Haematopoietic cell transfection enables gene therapy, gene
 CC vaccination and anti-sense therapy of diseases involving haematopoietic
 CC cells, including leukaemia and bone marrow stem cell disorders.
 CC Transfection of corneal endothelium is useful for treatment of eye
 CC disease affecting the cornea or corneal organ transplants, for e.g. in
 CC glaucoma. A gene preventing cell proliferation in blood vessel walls is
 CC introduced using an integrin targeting transfection vector complex (III)
 CC to reduce restenosis. (III) is useful for intracellular transport and
 CC delivery of anti-sense oligonucleotides, which enables antiviral and
 CC cancer therapy and is effective in transporting large DNA molecules.
 CC This sequence represents a cyclic peptide that lacks the conserved RGD
 CC integrin binding amino acid sequence but binds to integrins to allow
 CC the nucleic acid to pass into the cell, described in the method of the
 CC invention.

SQ

Sequence 31 AA:

Query Match 80.6%; Score 50; DB 23; Length 31;
 Best Local Similarity 75.0%; Pred. No. 0.093; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAXXXETAWACG 12
 Db 20 GACRRETAWACG 31

RESULT 15

AAU74982 ID AAU74982 standard; Peptide: 10 AA.

AC AAU74982;
 XX

DT 09-APR-2002 (first entry)

DE Transfection associated, integrin binding peptide #6.

XX Cyclic; virucide; human immunodeficiency virus; HIV; cytostatic;
 KW ophthalmological; vasotropic; vaccine; gene therapy; transfection;
 KW cystic fibrosis; asthma; cancer; leukaemia; glaucoma; gene vaccination;
 KW anti-sense therapy; eye disease; corneal organ transplant; integrin;
 KW transfection; restenosis.

OS Synthetic.

XK PH key Location/qualifiers

PT Disulphide_bond 1..9
 PT /note= "Cysteine residues linked by a disulfide
 bond to form a cyclic structure".

PN WO200192543-A2.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-GB02396.

XX PR 30-MAY-2000; 2000GB-0013099.

PR 01-MAY-2001; 2001US-287410P.

XX PA (ICHI-) ICH PRODN LTD.

PT Hart SL;

XX XPL WPI; 2002-114355/15.

XX
 PT Transflecting confluent cells with nucleic acid for gene therapy or gene
 PT vaccination, comprises contacting the cells with a receptor-targeted
 PT vector having the nucleic acid and an agent that disrupts cell-cell
 PT junctions

XX PS Claim 17; Page 17; 11pp; English.

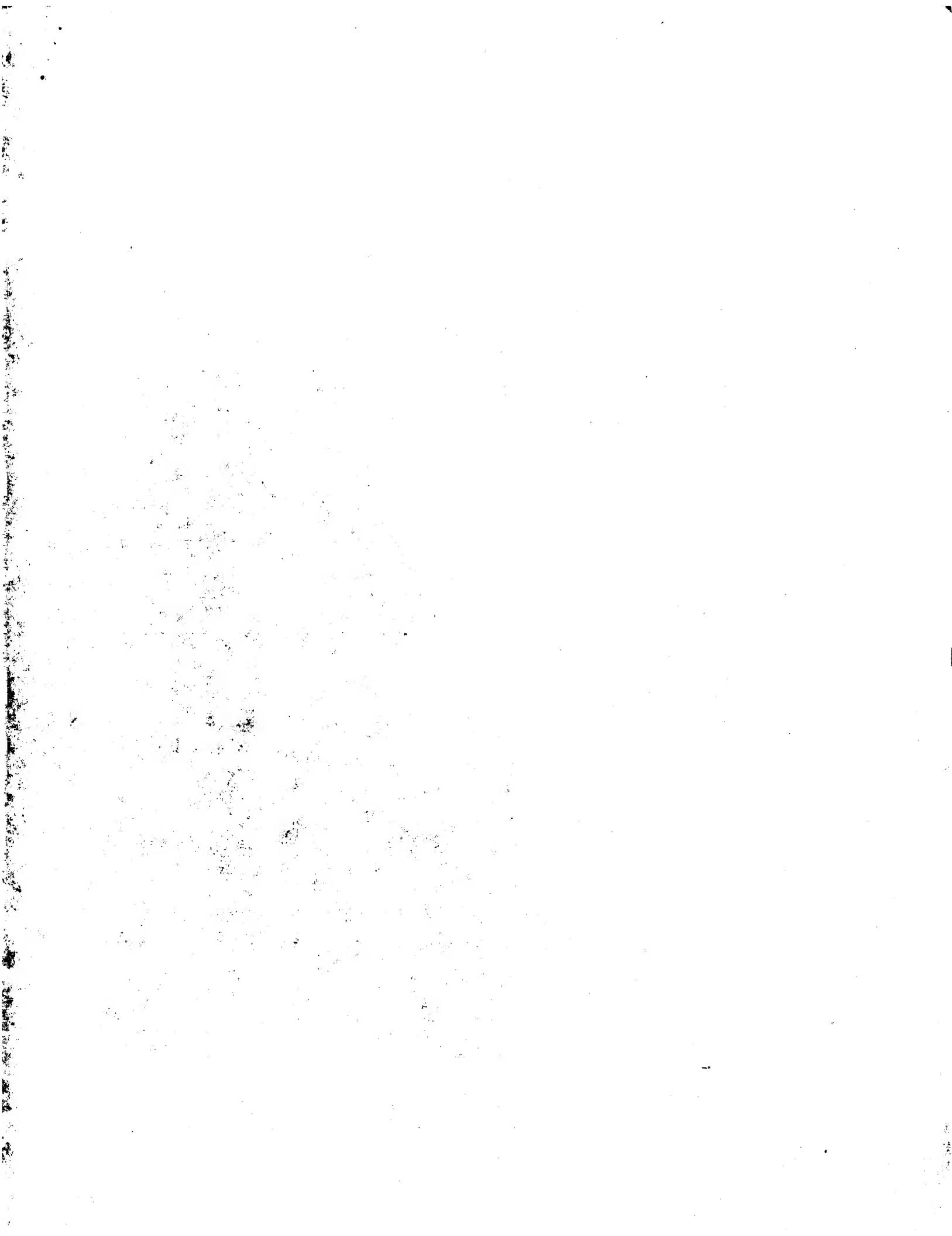
XX
 CC The invention describes transfecting (I) confluent cells or other slowly
 CC dividing or non-dividing cells that are in contact with each other, with
 CC a nucleic acid. The method comprises contacting the cells with a
 CC receptor-targeted vector comprising the nucleic acid, and an agent that
 CC disrupts cell-cell junctions under conditions suitable to effect
 CC transfection. (II) is useful for transfecting bronchial and lung
 CC epithelium for gene therapy for cystic fibrosis, asthma and also various
 CC cancers and viral infections e.g. human immunodeficiency virus (HIV)
 CC infection. Haematopoietic cell transfection enables gene therapy, gene
 CC vaccination and anti-sense therapy of diseases involving haematopoietic
 CC cells, including leukaemia and bone marrow stem cell disorders.
 CC Transfection of corneal endothelium is useful for treatment of eye
 CC disease affecting the cornea or corneal organ transplants, for e.g. in
 CC glaucoma. A gene preventing cell proliferation in blood vessel walls is
 CC introduced using an integrin targeting transfection vector complex (III)
 CC to reduce restenosis. (III) is useful for intracellular transport and
 CC delivery of anti-sense oligonucleotides, which enables antiviral and
 CC cancer therapy and is effective in transporting large DNA molecules.
 CC This sequence represents a peptide that will permit cyclisation by
 CC disulfide bond formation. It lacks the conserved RGD amino acid
 CC integrin binding sequence but binds integrins, allowing the nucleic
 CC acid to pass into the cell, described in the method of the invention.

XX SQ Sequence 10 AA:

Query Match 71.0%; Score 44; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.33; Mismatches 0; Indels 0; Gaps 0;

QY 6 ETAWACG 12
 Db 4 ETAWACG 10

Search completed: March 25, 2003, 08:19:41
 Job time : 41.7273 secs



Gencore version 5.1.4_p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - Protein search, using sw model

Run on: March 25, 2003, 08:21:42 ; Search time 14.8485 Seconds

27.742 Million cell updates/sec

Title: US-09-646-532B-2

Perfect score: 62

Sequence: 1 GAXXETAWACGXA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents,AAs:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCUTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	80.6	12	4 US-09-424-656-12	Sequence 12, Appl
2	50	80.6	13	1 US-08-425-238-3	Sequence 3, Appl
3	50	80.6	13	2 US-08-286-861-6	Sequence 6, Appl
4	50	80.6	13	4 US-09-406-781-63	Sequence 63, Appl
5	50	80.6	25	4 US-09-424-656-16	Sequence 16, Appl
6	44	71.0	12	4 US-09-424-656-11	Sequence 11, Appl
7	40	64.5	13	13 US-08-425-238-11	Sequence 11, Appl
8	38	61.3	9	1 US-08-425-238-6	Sequence 6, Appl
9	38	61.3	9	2 US-08-717-169-18	Sequence 18, Appl
10	38	61.3	9	2 US-08-286-12	Sequence 12, Appl
11	38	61.3	9	4 US-09-424-656-13	Sequence 13, Appl
12	38	61.3	13	2 US-09-701-124-35	Sequence 35, Appl
13	38	61.3	13	3 US-09-225-35	Sequence 35, Appl
14	38	61.3	13	4 US-09-455-061-35	Sequence 35, Appl
15	35	56.5	237	4 US-09-111-470-2	Sequence 2, Appl
16	35	56.5	662	4 US-08-79-814-5	Sequence 5, Appl
17	35	56.5	928	3 US-09-320-878-13	Sequence 13, Appl
18	35	56.5	928	4 US-09-105-537-41	Sequence 41, Appl
19	35	56.5	1130	2 US-09-519-547-A6	Sequence 6, Appl
20	35	56.5	3155	2 US-08-459-146-3	Sequence 3, Appl
21	35	56.5	3165	2 US-08-459-065-3	Sequence 2, Appl
22	34	54.8	222	3 US-08-651-136C-14	Sequence 14, Appl
23	34	54.8	222	4 US-09-229-911A-14	Sequence 14, Appl
24	34	54.8	294	3 US-08-651-136C-24	Sequence 24, Appl
25	34	54.8	294	4 US-09-229-911A-24	Sequence 24, Appl
26	33	53.2	170	4 US-09-195-637A-299	Sequence 299, Appl
33	53.2	53.2	376	4 US-09-135-010A-113	Sequence 113, App
34	53.2	53.2	376	4 US-09-444-871-113	Sequence 113, App
35	53.2	53.2	376	4 US-09-444-295-113	Sequence 113, App
36	53.2	53.2	376	4 US-09-597-732-113	Sequence 113, App
37	53.2	53.2	376	4 US-09-444-871-114	Sequence 114, App
38	53.2	53.2	570	4 US-09-597-735-114	Sequence 114, App
39	53.2	53.2	570	4 US-09-444-995-114	Sequence 114, App
40	53.2	53.2	570	4 US-09-597-732-114	Sequence 114, App
41	53.2	53.2	581	4 US-09-135-021-B80	Sequence 116, App
42	53.2	53.2	581	4 US-09-135-010-A16	Sequence 116, App
43	53.2	53.2	581	4 US-09-597-735-116	Sequence 116, App
44	53.2	53.2	581	4 US-09-597-732-116	Sequence 116, App
45	53.2	53.2	605	4 US-09-105-058C-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
 US-09-424-656-12
 Sequence 12, Application US/09424656
 ; Patent No. 6458026
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: ENHANCED TRANSFECTION ACTIVITY
 NUMBER OF SEQUENCES: 16
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/424, 656
 FILING DATE:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: GB 9711115.7
 FILING DATE: 29-MAY-1997
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12 amino acids
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-424-656-12
 Query Match Best Local Similarity 80.6%; Score 50; DB 4; Length 12;
 Matches 9; Conservative 75.0%; Pred. No. 0.011; 3; Indels 0; Gaps 0;
 Qy 1 GAXXETAWACG 12
 Db 1 GACRRETAWACG 12

RESULT 2
 US-08-425-238-3
 Sequence 3, Application US/08425238
 ; Patent No. 5027263
 GENERAL INFORMATION:
 APPLICANT: Ruuslahti, EERKI
 APPLICANT: KOIVILAINEN, EERKI
 TITLE OF INVENTION: NO. 5627263el Integrin-Binding Peptides
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 Sequence 113, App

CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425, 238
FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/158, 001
FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31, 815
REFERENCE/DOCKET NUMBER: P-LA 9775

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-425-238-3

Query Match 80 6%; Score 50; DB 1; Length 13;
Best Local Similarity 75 0%; Pred. No. 0.012; Mismatches 0;
Matches 9; Conservative 0; Indels 3; Gaps 0;
QY 1 GAXXXETAWACG 12
Db 1 GACRRETAWACG 12

RESULT 3

US-08-286-861-6
Sequence 5: Application US/08286861
Patent No. 598178

GENERAL INFORMATION:

APPLICANT: Ruoslahti, Erkki

APPLICANT: Koivunen, Erkki

TITLE OF INVENTION: No. 5981478el Integrin-Binding Peptides

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

STATE: California

COUNTY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286, 861
FILING DATE: 04-AUG-1994
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/158, 001
FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31, 815

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

RESULT 4

US 09-406-781-63
Sequence 63: Application US/09406781
Patent No. 630663

GENERAL INFORMATION:

APPLICANT: Kenten, John

APPLICANT: Roberts, Steven

TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS

FILE REFERENCE: 2757-3

CURRENT APPLICATION NUMBER: US/09/406, 781

CURRENT FILING DATE: 1999-09-28

EARLIER APPLICATION NUMBER: 60/119, 851

EARLIER FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 63

LENGTH: 13

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE:

OTHER INFORMATION: Description of unknown Organism: binding peptide

US-09-406-781-63

Query Match 80 6%; Score 50; DB 4; Length 13;
Best Local Similarity 75 0%; Pred. No. 0.012; Mismatches 0;
Matches 9; Conservative 0; Indels 3; Gaps 0;

QY 1 GAXXXETAWACG 12

Db 1 GACRRETAWACG 12

RESULT 5

US-09-424-656-16
Sequence 16: Application US/09424656
Patent No. 6458026

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING ENHANCED TRANSEFCTION ACTIVITY

NUMBER OF SEQUENCES: 16

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/424, 656

FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9711115.7

FILING DATE: 29-MAY-1997

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: circular
 MOLECULE TYPE: peptide

Query Match 80.5%; Score 50; DB 4; Length 25;
 Best Local Similarity 75.0%; Pred. No. 0.022; Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAXXXETAWACCG 12
 Db 1 GAXRRETAWACG 12

RESULT 6
 US-09-424-656-11
 Sequence 11, Application US/09424656
 Patent No. 6,58026

GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING ENHANCED TRANSFECTION ACTIVITY
 TITLE OF INVENTION: ENHANCED TRANSFECTION ACTIVITY
 NUMBER OF SEQUENCES: 16
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/424,656
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9711115.7
 FILING DATE: 29-MAY-1997

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
 LENGTH: 12 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: circular
 MOLECULE TYPE: Peptide

US-09-424-656-11

RESULT 7
 Query Match 71.0%; Score 44; DB 4; Length 12;
 Best Local Similarity 72.7%; Pred. No. 0.11; Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAXXXETAWAC 11
 Db 1 GAXRRETAWAC 11

RESULT 7
 US-08-425-238-11
 Sequence 11, Application US/08425238
 Patent No. 5,627263

GENERAL INFORMATION:
 APPLICANT: Ruoslahti, Erika
 APPLICANT: Koivinen, Erkki
 TITLE OF INVENTION: No. 5627263el Integrin-Binding Peptides
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/425,238
 FILING DATE:
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/158,001
 FILING DATE: 24-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LA 9775

SEQUENCE CHARACTERISTICS:
 LENGTH: 13 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:
 NAME/KEY: Peptide
 LOCATION: 3
 OTHER INFORMATION: /note= "Xaa = an amino acid capable of forming a disulfide bond."

FEATURE:
 NAME/KEY: Peptide
 LOCATION: 11
 OTHER INFORMATION: /note= "Xaa = an amino acid capable of forming a disulfide bond."

OTHER INFORMATION: Of forming a disulfide bond.

US-08-425-238-11

RESULT 8
 Query Match 64.5%; Score 40; DB 1; Length 13;
 Best Local Similarity 75.0%; Pred. No. 0.56; Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAXXXETAWACG 12
 Db 1 GAXRRETAWACG 12

RESULT 8
 US-08-425-238-6
 Sequence 6, Application US/08425238
 Patent No. 5,627263

GENERAL INFORMATION:
 APPLICANT: Ruoslahti, Erika
 APPLICANT: Ruoslahti, Erika
 APPLICANT: Koivinen, Erkki
 TITLE OF INVENTION: No. 5627263el Integrin-Binding Peptides
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/425,238
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LA 9775
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
; US-08-425-238-6

Query Match 61.3%; Score 38; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+05; Length 9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 ETAWAC 11
 Db 4 ETAWAC 9

RESULT 9
; US-08-717-169-18
; Sequence 18, Application US/08717169
; GENERAL INFORMATION:
; APPLICANT: Pasqualini, Renata
; APPLICANT: Ruošlahti, Erkki
; TITLE OF INVENTION: Methods of Inhibiting Angiogenesis and
; TITLE OF INVENTION: Ameliorating Cancer By Using Superfibronectin
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 REFERENCE/DOCKET NUMBER: P-LA 9775
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/158,001
 FILING DATE: 24-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LA 9992
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: circular
; US-08-286-861-12

Query Match 61.3%; Score 38; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+05; Length 9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 ETAWAC 11
 Db 4 ETAWAC 9

RESULT 11
; US-09-424-656-13
; Sequence 13, Application US/09424656
; Patent No. 6458026
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING
; TITLE OF INVENTION: ENHANCED TRANSFECTION ACTIVITY
; NUMBER OF SEQUENCES: 16
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO),
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/424,656
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 971115.7
; FILING DATE: 29-MAY-1997
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids

RESULT 10

TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: circular
 MOLECULE TYPE: Peptide
 US-09-424-656-13

Query Match 61.3%; Score 38; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Qy 6 ETAWAC 11
 Db 4 ETAWAC 9

RESULT 12
 US-08-701-124-35
 ; Sequence 35, Application US/08701124
 ; Patent No. 5846782
 GENERAL INFORMATION:
 APPLICANT: Wickham, Thomas J.
 APPLICANT: Roelvink, Petrus W.
 APPLICANT: Kovendi, Imre

TITLE OF INVENTION: TARGETING ADENOVIRUS WITH USE OF
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 STREET: Two Prudential Plaza - 49th Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08-701-124
 FILING DATE: 21-AUG-1996
 INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:
 LENGTH: 13 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide

RESULT 13
 US-08-701-124-35
 Query Match 61.3%; Score 38; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

Qy 6 ETAWAC 11
 Db 6 ETAWAC 11

RESULT 14
 US-09-455-061-35
 ; Sequence 35, Application US/09455061
 ; Patent No. 6329190
 GENERAL INFORMATION:
 APPLICANT: Wickham, Thomas J.
 APPLICANT: Roelvink, Petrus W.
 APPLICANT: Kovendi, Imre

TITLE OF INVENTION: TARGETING ADENOVIRUS WITH USE OF
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 STREET: Two Prudential Plaza - 49th Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/455, 061
 FILING DATE: 06-DEC-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 9-130225
 FILING DATE: 06-AUG-1998

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 8-701124
 FILING DATE: 21-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Heffner, M. Daniel
 REGISTRATION NUMBER: 41,826
 REFERENCE/DOCKET NUMBER: 203128

INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 13 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide

US-09-455-061-35

Query Match Similarity 61.3%; Score 38; DB 4; length 13;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6 ETAWAC 11 	Applicant: Ford, John
Db	6 ETAWAC 11	Applicant: Saeland, Sem

RESULT 15

US-09-111-470-2

Sequence 2, Application US/09111470

Patent No. 6277959

GENERAL INFORMATION:

Applicant: Valladau, Jenny

Applicant: Ravel, Odile

Applicant: Bates, Elizabeth E.M.

Applicant: Lebecque, Serge J.E.

Title of Invention: Mammalian Membrane Protein Genes; Related Reagents

Number of Sequences: 11

Correspondence Address:

Addressee: DNAX Research Institute

Street: 901 California Avenue

City: Palo Alto

State: California

Country: USA

ZIP: 94304-1104

Computer Readable Form:

Medium Type: Floppy disk

Computer: IBM PC compatible

Operating System: PC-DOS/MS-DOS

Software: PatentIn Release #1.0, Version #1.30

Current Application Data:

Application Number: US/09/111,470

Filing Date: 08-JUL-1998

Classification:

Prior Application Data:

Application Number: US 60/053, 080

Filing Date: 09-JUL-1997

Attorney/Agent Information:

Name: Ching, Edw-in P.

Registration Number: 34,090

Reference/Docket Number: SF0695

Telecommunication Information:

Telephone: (650)852-9196

Telex/Fax: (650)96-1200

Information for Seq ID No: 2:

Sequence Characteristics:

Length: 237 amino acids

Type: amino acid

Topology: linear

Molecule Type: protein

US-09-111-470-2

Query Match Similarity 56.5%; Score 35; DB 4; Length 237;
 Best Local Similarity 83.3%; Pred. No. 61; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	6 ETAWAC 11
Db	100 ETANSC 105

Search completed: March 25, 2003, 08:22:59
 Job time : 15.8485 secs

GenCore version 5.1.4_P5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 08:22:27 ; Search time 13.1515 seconds
 (without alignments)
 56.911 Million cell updates/sec

Title: US-09-646-532B-2

Perfect score: 62

Sequence: 1 GAXXETAWACGXA, 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_AA:**

1: /cgn2_6/prodata/2/pubpaas/PCT_NEW_PUB.pep:**
 2: /cgn2_6/prodata/2/pubpaas/PCT_NEW_PUB.pep:**
 3: /cgn2_6/prodata/2/pubpaas/US06_PUBCOMB.pep:**
 4: /cgn2_6/prodata/2/pubpaas/US07_NEW_PUB.pep:**
 5: /cgn2_6/prodata/2/pubpaas/US07_PUBCOMB.pep:**
 6: /cgn2_6/prodata/2/pubpaas/US07_PUBCOMB.pep:**
 7: /cgn2_6/prodata/2/pubpaas/PCTUS_PUBCOMB.pep:**
 8: /cgn2_6/prodata/2/pubpaas/US08_PUBCOMB.pep:**
 9: /cgn2_6/prodata/2/pubpaas/US09_NEW_PUB.pep:**
 10: /cgn2_6/prodata/2/pubpaas/US09_PUBCOMB.pep:**
 11: /cgn2_6/prodata/2/pubpaas/US10_NEW_PUB.pep:**
 12: /cgn2_6/prodata/2/pubpaas/US10_PUBCOMB.pep:**
 13: /cgn2_6/prodata/2/pubpaas/US60_NEW_PUB.pep:**
 14: /cgn2_6/prodata/2/pubpaas/US60_PUBCOMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	50	80	6	12 9 US-10-137-435-12 Sequence 12, Appl
2	50	80	6	13 9 US-09-880-132-63 Sequence 63, Appl
3	50	80	6	13 10 US-09-364-597A-6 Sequence 6, Appl
4	50	80	6	13 10 US-09-880-149-63 Sequence 63, Appl
5	50	80	6	13 10 US-10-137-435-16 Sequence 16, Appl
6	44	71	0	12 9 US-10-137-435-11 Sequence 11, Appl
7	38	61	3	9 US-10-137-435-13 Sequence 13, Appl
8	38	61	3	9 10 US-09-364-597A-12 Sequence 12, Appl
9	38	61	3	13 10 US-09-969-192-35 Sequence 35, Appl
10	37	59	7	38 10 US-09-864-761-38378 Sequence 38378_A
11	36	58	1	158 9 US-09-883-152-4 Sequence 4, Appl
12	36	58	1	166 10 US-09-747-835A-56 Sequence 56, Appl
13	36	58	1	310 10 US-09-886-055-495 Sequence 485, App
14	36	58	1	2213 9 US-10-188-64-549 Sequence 549, App
15	35	56	5	202 10 US-09-280-197-11 Sequence 11, Appl
16	35	56	5	237 9 US-09-862-892-9 Sequence 2, Appl
17	35	56	5	297 9 US-09-729-49 Sequence 49, Appl
18	35	56	5	243 10 US-09-764-870-464 Sequence 464, App
19	35	56	5	246 10 US-09-764-870-303 Sequence 303, App

ALIGNMENTS

RESULT 1
 US-10-137-435-12
 ; Sequence 12, Application US/10137435
 ; Publication No. US20030011364A1
 ; GENERAL INFORMATION:
 APPLICANT: Institute of Child Health
 TITLE OF INVENTION: ENHANCED TRANSEFCTION VECTORS HAVING
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 STREET: c/o Institute of Child Health,
 30 Guildford Street
 CITY: London
 COUNTRY: G.B.
 ZIP: WC1N 1EH

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/137,435
 FILING DATE: 03-May->2002
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/09/424,656
 FILING DATE: <Unknown>
 APPLICATION NUMBER: GB 9711115.7
 FILING DATE: 29-MAY-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Hart, Stephen Lewis
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-137-435-12

Query Match Best Local Similarity 80.6%; Score 50; DB 9; Length 12;

Best Local Similarity 75.0%; Pred. No. 0.014;

Matches 9; conservative 0; Mismatches 3; Indels 0; Gaps 0;

NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
TELECOMMUNICATION NUMBER: P-LA 3419
TELEPHONE: (858) 535-9001
TELEFAX: (858) 535-8949

RESULT 2
US-09-880-132-63

; Sequence 63, Application US/09880132
; Patent No. US20017049A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John
; ROBERTS, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 275-6
; CURRENT APPLICATION NUMBER: US/09-880,132
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/406,781
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/119,851
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: binding peptide

Query Match 80.6%; Score 50; DB 9; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

US-09-880-132-63

RESULT 3
US-09-364-597A-6

; Sequence 6, Application US/09364597A
; Patent No. US20020103130A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Koivunen, Erkki
; TITLE OF INVENTION: No. US20020103130A1 Hel Integrin-Binding Peptides
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/364,597A
; FILING DATE: 30-JUL-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/158, 001
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286, 861
; FILING DATE: 04-AUG-1994
; ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
TELECOMMUNICATION NUMBER: P-LA 3419
TELEPHONE: (858) 535-9001
TELEFAX: (858) 535-8949

RESULT 4
US-09-880-149-63

; Sequence 63, Application US/09880149
; Patent No. US20020146843A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John
; ROBERTS, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 275-5
; CURRENT APPLICATION NUMBER: US/09-880,149
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/406,781
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/119,851
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: binding peptide

US-09-880-149-63

RESULT 5
US-10-137-435-16

; Sequence 16, Application US/10137435
; Publication No. US20030013644A1
; GENERAL INFORMATION:
; APPLICANT: Institute of Child Health
; TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING ENHANCED TRANSEFCTION ACTIVITY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: c/o Institute of Child Health,
; CITY: London
; COUNTRY: G.B.
; ZIP: WC1N 1EH
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/424,656

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Hart, Stephen Lewis

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: circular

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-10-137-435-16

Query Match 80.6% Score 50; DB 9; Length 25;
Best Local Similarity 75.0%; Pred. No. 0.027; Mismatches 0;
Matches 9; Conservative 0; Indels 0; Gaps 0;

QY 1 GAXXXETAWAC 12
Db 1 GACRRETAWAC 12

RESULT 6
US-10-137-435-11

Sequence 11, Application US/10137435
Publication No. US2003013644A1

GENERAL INFORMATION:

APPLICANT: Institute of Child Health

TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING ENHANCED TRANSFECTION ACTIVITY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

STREET: c/o Institute of Child Health,
30 Guildford Street
CITY: London
COUNTRY: G.B.
ZIP: WC1N 1EH

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.30 (EPO)

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/10/137,435

FILING DATE: 03-May-2002

ATTORNEY/AGENT INFORMATION:

NAME: Hart, Stephen Lewis

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: circular

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-137-435-13

Query Match 61.3% Score 38; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05; Mismatches 0;
Matches 6; Conservative 0; Indels 0; Gaps 0;

QY 6 ETAWAC 11
Db 4 ETAWAC 9

RESULT 7
US-10-137-435-13

Sequence 13, Application US/10137435
Publication No. US2003013644A1

GENERAL INFORMATION:

APPLICANT: Institute of Child Health

TITLE OF INVENTION: INTEGRIN-TARGETING VECTORS HAVING ENHANCED TRANSFECTION ACTIVITY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

STREET: c/o Institute of Child Health,
30 Guildford Street
CITY: London
COUNTRY: G.B.
ZIP: WC1N 1EH

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.30 (EPO)

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/424,656

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Hart, Stephen Lewis

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: circular

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-137-435-13

Query Match 61.3% Score 38; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05; Mismatches 0;
Matches 6; Conservative 0; Indels 0; Gaps 0;

QY 6 ETAWAC 11
Db 4 ETAWAC 9

RESULT 8
US-09-364-597A-12

Sequence 12, Application US/09364597A
Patent No. US20020103130A1

GENERAL INFORMATION:

APPLICANT: Ruoslahti, Erkki

APPLICANT: Koivunen, Erkki

TITLE OF INVENTION: No. US20020103130A1 Integrin-Binding Peptides

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESS: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92123

US-10-137-435-11

Query Match 71.0% Score 44; DB 9; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.15;
Matches 8; Conservative 0; Mismatches 3;
Indels 0; Gaps 0;

QY 1 GAXXXETAWAC 11
Db 1 GACRRETAWAC 11

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/364,597A
FILING DATE: 30-JUL-1999
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/158,001
FILING DATE: 24-NOV-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/286,861
FILING DATE: 04-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-PLA 3419

TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 535-9001
TELEFAX: (856) 535-8949

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
MOLECULE TYPE: Peptide
TOPOLogy: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 35

US-09-969-192-35

Query Match 61.3%; Score 38; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ETAWAC 11
Db 6 ETAWAC 11

RESULT 9
US-09-969-192-35
Sequence 35, Application US/09959192
; Patent No. US20020151027A1
; GENERAL INFORMATION.
APPLICANT: WICKHAM, THOMAS J.
ROEVINK, IMRE
KOVESDI, IMRE
TITLE OF INVENTION: TARGETING ADENOVIRUS WITH USE OF
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza - 49th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/969,192
FILING DATE: 01-OCT-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 9-455051
FILING DATE: 06-DEC-1999
APPLICATION NUMBER: US 9-130225
FILING DATE: 06-AUG-1998
APPLICATION NUMBER: US 8-701124
FILING DATE: 21-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: Hefner, M. Daniel

RESULT 10
US-09-864-761-38378
Sequence 38378, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 242633.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Animax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38378

Page
5

LENGTH: 38
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC005630.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.2
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.9
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.7
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.3
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.5
 OTHER INFORMATION: OTHER INFORMATION: EST_HUMAN HIT: AW162304.1, EVALUE 3.00e-05
 US-09-864-761-38378

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Indels	Gaps
QY 1 GAXXXETAVACGXA 14	59.7%	37	DB 10	Length 38;			0	0;
QY 1 111 37	50.0%	5.9						
Db 24 GAPVPOGAWMKCGQA 37								

RESULT 11
 US-09-883-152-4
 Sequence 4, Application US/09883152
 Publication No. US20030008284A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kennedy, Giulia
 ; APPLICANT: Kang, Sammo
 ; APPLICANT: Reinhard, Christoph
 ; APPLICANT: Jefferson, Anne Bennett
 ; TITLE OF INVENTION: POLYNUCLEOTIDES RELATED TO COLON CANCER
 ; FILE REFERENCE: 2300-1663
 ; CURRENT APPLICATION NUMBER: US/09/883,152
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/211,835
 ; PRIOR FILING DATE: 2000-06-15
 ; NUMBER OF SEQ ID NOS: 127
 ; SOFTWARE: FastaSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 158
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 ; US-09-883-152-4

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Indels	Gaps
QY 7 TAWACG 12	59.7%	37	DB 10	Length 38;			0	0;
QY 7 TAWACG 12	50.0%	5.9						
Db 30 TSWACG 35								

RESULT 12
 US-09-747-835A-56
 Sequence 56, Application US/09747835A
 Patent No. US20020146692A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamazaki, Victoria
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyun
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Dmircac, Radivoj T
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
 ; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
 ; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Indels	Gaps
QY 1 GAXXXETAVACG 12	58.1%	36	DB 9	Length 158;			0	0;
QY 1 111 108	41.7%	29						
Db 97 GQTREDLANSCG 108								

RESULT 13
 US-09-886-055-485
 Sequence 485, Application US/09886055
 Patent No. US20030132273A1
 ; GENERAL INFORMATION:
 ; APPLICANT: STRYER, LOBERT
 ; APPLICANT: ZOZULYA, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; FILE REFERENCE: 078003-027150
 ; CURRENT APPLICATION NUMBER: US/09/886, 055
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213, 812
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 485
 ; LENGTH: 310
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-886-055-485

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Indels	Gaps
QY 7 TAWACG 12	58.1%	36	DB 10	Length 310;			0	0;
QY 7 TAWACG 12	50.0%	5.2						
Db 146 TSWACG 151								

RESULT 14
 US-10-184-644-549
 Sequence 549, Application US/10184644
 Publication No. US20030044030A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watnabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430RIC227
 CURRENT APPLICATION NUMBER: US/10/184,644
 CURRENT FILING DATE: 2002-06-28
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO: 549
 LENGTH: 2213
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-184-644-549

Query Match 58.1%; Score 36; DB 9; Length 2213;
 Best Local Similarity 57.1%; Pred. No. 2.8e+02;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GAXXETAWAGCGXA 14
 Db 1752 GAGTCATACACGAA 1765

RESULT 15
 US-09-280-197-11
 Sequence 11, Application US/09280197
 Patent No. US2002014203A1
 GENERAL INFORMATION:
 APPLICANT: Yu, Shikun
 APPLICANT: Bojsen, Kirsten
 APPLICANT: Kragh, Karsten
 APPLICANT: Bojko, Maja
 APPLICANT: Nielsen, John
 APPLICANT: Marcusen, Jan
 APPLICANT: Christensen, Tove
 TITLE OF INVENTION: USE OF 1,4-GLUCAN LYASE FOR PREPARATION OF 1,5-D-ANHYDROFRUCTOSE
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe, Martens, Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/280,197
 FILING DATE: OCT-15-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/633,719
 FILING DATE: July 8, 1996
 APPLICATION NUMBER: PCT/EP94/03397
 FILING DATE: OCT-15-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Altman, Daniel E
 REGISTRATION NUMBER: 34,115
 REFERENCE/DOCKET NUMBER: DYOUS.001APC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714-750-0404
 TELEFAX: 714-760-9502
 TELEX:
 INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
 LENGTH: 202 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 43
 OTHER INFORMATION: /note= "X is a misc. amino acid"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 176
 OTHER INFORMATION: /note= "X is a misc. amino acid"
 US-09-280-197-11
 Query Match 56.5%; Score 35; DB 10; Length 202;
 Best Local Similarity 71.4%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 6 ETAWAGCG 12
 Db 68 DTAWNCG 74

Search completed: March 25, 2003, 08:23:40
 Job time : 15.1515 secs

Query Match 74.2%; Score 46; DB 2; Length 290;
 Best Local Similarity 64.3%; Pred. No. 0.53;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GAXXXETAWACGXA 14
 Db 251 GASLAAETAFACGFA 264

RESULT 3
 A83271 hypothetical protein PA984 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: A83271
 R;Slover, C.K.; Pham, X.Q.; Erwin, A.L.; Mioguchi, S.D.; Warrener, P.; Hickey, M.J.; Braden, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Mory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: A83271
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-741 <HII>
 A;Cross-references: GB:AE004724; GB:AE004091; NID:99949083; PIDN:AAG06372.1; GSPDB:GN001
 A;Experimental source: strain PA01
 C;Genetics:

A;Gene: PA2984

Query Match 61.3%; Score 38; DB 2; Length 741;
 Best Local Similarity 85.7%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 8 AWAGCGXA 14
 Db 53 AWACGSA 59

RESULT 4
 G87548 transglycosylase, probable CC2416 [imported] - *Caulobacter crescentus*
 C;Species: *Caulobacter crescentus*
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C;Accession: G87548
 R;Neerman, W.C.; Heidbrink, T.V.; Paulsen, I.T.; Nelson, K.E.; Bisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Halt, D.H.; Kolon, N.; U.; Emelava, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: G87548
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-218 <STO>
 A;Cross-references: GB:AE005673; NID:913423957; PIDN:AAK24387.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC2416

Query Match 59.7%; Score 37; DB 2; Length 218;
 Best Local Similarity 85.7%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 8 AWAGCGXA 14
 Db 22 AWACGVA 28

RESULT 5
 ET7548 Query Match 59.7%; Score 37; DB 2; Length 334;
 Best Local Similarity 59.7%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 GAXXXETAWACGXA 14
 Db 115 GASGTTELAAACGSA 128

RESULT 6
 T34634 probable hydrolase - *Streptomyces coelicolor*
 C;Species: *Streptomyces coelicolor*
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C;Accession: T34634
 R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, A;Reference number: Z21549
 A;Accession: T34634
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-334 <SAU>
 A;Cross-references: EMBL:AL078618; PIDN:CAB44541.1; GSPDB:GN00070; SCOEDB:SC10A7.31
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SC10A7.31

Query Match 59.7%; Score 37; DB 2; Length 334;
 Best Local Similarity 57.1%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 GAXXXETAWACGXA 14
 Db 115 GASGTTELAAACGSA 128

RESULT 7
 S73770 NADH oxidase nox - *Mycoplasma pneumoniae* (strain ATCC 29342)
 N;Alternate names: hypothetical protein Fil_orf479
 C;Species: *Mycoplasma pneumoniae*
 A;Variety: ATCC 29342
 C;Accession: S73770
 C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkle, E.; Li, B.C.; Herrmann, R.; Nucleic Acids Res. 24, 4420-4449, 1996
 A;Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*
 A;Reference number: S73327; MUID:97105885; PMID:8948633
 A;Accession: S73770
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-479 <HII>
 A;Cross-references: EMBL:AE00044; GB:U00089; NID:91674130; PIDN:AAB96092.1; PID:9167
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C;Genetics:

A;Genetic code: SGC3
 C;Superfamily: NADH peroxidase; dihydrolipamide dehydrogenase homology
 Query Match 59.7%; Score 37; DB 2; Length 479;
 C;Accession: C09999
 Best Local Similarity 50.0%; Pred. No. 36; Mismatches 6; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 GAXXXETAWACG 12
 Db 179 GLELAEAAWQCG 190

RESULT 8
 F85715 hypothetical protein Z2116 [imported] - Escherichia coli (strain 0157:H7, substrate EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: F85715
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: F85715
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-502 <STO>
 A;Cross-references: GB:AE005174; NID:912515071; PIDN:AAG56186.1; GSPDB:GN00145; UWGP:z21
 A;Experimental source: strain 0157:H7, substrate EDL933
 C;Genetics:
 C;Gene: 22116

Query Match	59.7%	Score	37;	DB	2;	Length	502;
Best Local Similarity	42.9%	Pred. No.	38;				
Matches	6;	Conservative	1;	Mismatches	7;	Indels	0;
Qy	1 GAXXXETANACGXA 14						
Db	178 GRDITDKWRCGGA 191						

RESULT 9
 A85819 hypothetical protein Z3099 [imported] - Escherichia coli (strain 0157:H7, substrate EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: A85819
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: A85819
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-157 <VAR>
 A;Cross-references: EMBL:X89663; NID:9002194; PIDN:CAA61810.1; PID:9002195
 C;Superfamily: olfactory receptor OR14
 Query Match 59.7%; Score 37; DB 2; Length 479;
 C;Accession: S88000
 Best Local Similarity 42.9%; Pred. No. 38; Mismatches 6; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 GAXXXETANACGXA 14
 Db 178 GRDITDKWRCGGA 191

RESULT 10
 C90999 probable terminase large subunit [imported] - Escherichia coli (strain 0157:H7, substrate EDL933)
 A;Accession: S22026
 A;Reference number: S22026

RESULT 11
 S88000 probable olfactory receptor tpcr71 - dog (fragment)
 C;Species: Canis lupus familiaris (dog)
 C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 03-Nov-2000
 C;Accession: S88000
 R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
 Submitted to the EMBL Data Library, July 1995
 A;Description: Male germ cells from several mammalian species express a specific repe
 A;Accession: S88000
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-157 <VAR>
 A;Cross-references: EMBL:X89663; NID:9002194; PIDN:CAA61810.1; PID:9002195
 C;Superfamily: olfactory receptor OR14
 Query Match 58.1%; Score 36; DB 2; Length 157;
 C;Accession: A6221
 Best Local Similarity 83.3%; Pred. No. 20; Mismatches 5; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 GAXXXETANACGXA 14
 Db 22 TSWAGC 27

RESULT 12
 A46221 abdominal segment formation protein pumilio - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 17-Oct-1997
 C;Accession: A6221; S22026
 R;Barber, D.B.; Wang, C.; Moore, J.; Dickinson, L.K.; Lehmann, R.
 Genes Dev. 6, 2312-2326, 1992
 A;Title: Pumilio is essential for function but not for distribution of the Drosophila
 A;Reference number: A46221; MUID:93093466; PMID:1459455
 A;Contents: embryo
 A;Accession: A6221
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-153 <BAR>
 A;Cross-references: GB:L07943; NID:9158190; PID:9158191
 A;Note: sequence extracted from NCBI backbone (NCBIN:120203, NCBIP:120204)
 R;Macdonald, P.M.
 Submitted to the EMBL Data Library, October 1991
 A;Reference number: S22026
 A;Accession: S22026

A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-361; 'A', 363-1102, 'R', 1104-1405, 'K', 1408-1495, 'V', 1497-1518, 'S', 1520-1533
 A; Cross-references: EMBL:X62589; MID:98393; PID:98394
 C; Genetics:
 A; Gene: FlyBase:pum
 A; Cross-references: FlyBase:FBgn003165
 Query Match Similarity 58.1%; Score 36; DB 2; Length 1533;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02; 1; Mismatches 1; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 6 ETAWACG 12
 Db 98 QTRWACG 104

RESULT 13
 JCT7608
 type II lectin-like immunoreceptor - human
 C; Species: Homo sapiens (man)
 C; Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C; Accession: JCT7608
 R; Huang, X.; Yuan, Z.; Chen, G.; Zhang, M.; Zhang, W.; Yu, Y.; Cao, X.
 Biochem. Biophys. Res. Commun. 281, 131-140, 2001
 A; Reference number: JCT7608; MUID:21092797; PMID:1178971
 A; Title: Cloning and characterization of a novel ITIM containing lectin-like immunoreceptor
 A; Content: Dendritic cells
 A; Accession: JCT7608
 A; Molecule type: mRNA
 A; Residues: 1-237 <HUA>
 A; Cross-references: GB:AF067800
 C; Comment: This receptor, highly homologous to the group of macrophage/hepatocyte lectins i
 cell, especially in migrating, antigen capturing and processing.
 C; Genetics:
 A; Gene: llr
 A; Map position: 12p13
 F; 45-69/Domain: transmembrane #status predicted <TMM>
 Query Match Similarity 56.5%; Score 35; DB 2; Length 237;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 6 ETAWAC 11
 Db 100 ETAWSC 105

RESULT 14
 AV9882
 histidine decarboxylase (EC 4.1.1.22) - human
 C; Species: Homo sapiens (man)
 C; Date: 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 16-Jun-2000
 C; Accession: A49882; S11492; S22623; A56625
 R; Yamamoto, K.; Ohtsu, H.; Tsuchikawa, M.; Higuchi, T.; Ishibashi, K.; Shida, A.; Shina
 J. Biol. Chem. 269, 1554-1559, 1994
 A; Title: Structure of the L-histidine decarboxylase gene.
 A; Reference number: A49882; MUID:94117478; PMID:8286622
 A; Accession: A49882
 A; Molecule type: DNA
 A; Residues: 1-662 <YAT>
 A; Cross-references: GB:D16593; NID:9516770; PIDN:BAA04015.1; PID:9516771
 A; Note: only intron-exon junctions shown
 A; Residues: 1-147, 'Q', 149-662 <YAM>
 A; Cross-references: EMBL:X54297; NID:932108; PIDN:CAA38196.1; PID:932109
 R; Mamune-Sato, R.; Yamauchi, K.; Tanno, Y.; Ohtsu, H.; Katayose, D.; Maeya
 Eur. J. Biochem. 209, 533-539, 1992

RESULT 15
 T31633
 hypothetical protein Y57A10A_k - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
 C; Accession: T31633
 R; Snyre, R.
 submitted to the EMBL Data Library, September 1999
 A; Reference number: 221048
 A; Accession: T31633
 A; Status: preliminary; translated from GB/EMBL/DDB/J
 A; Molecule type: DNA
 A; Residues: 1-668 <WIL>
 A; Cross-References: EMBL:AL117195; PIDN:CAB55016.1; CESP:Y57A10A_k
 A; Experimental source: clone Y57A10A
 C; Genetics:
 A; Gene: CESP:Y57A10A_k
 A; Introns: 22/3; 182/3; 219/2; 340/3; 385/1; 451/3
 C; Superfamily: Caenorhabditis elegans hypothetical protein Y57A10A_k
 Query Match Similarity 56.5%; Score 35; DB 2; Length 668;
 Best Local Similarity 45.5%; Pred. No. 1.1e+02; 5; Mismatches 5; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 GAXXXETAWAC 11
 Db 326 GSTAHTGWAC 336

Search completed: March 25, 2003, 08:22:19
 Job time : 18.277 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 25, 2003, 08:18:08 ; Search time 8.48495 Seconds

(without alignments)

68.436 Million cell updates/sec

Title: US-09-646-532B-2

Perfect score: 62

Sequence: 1 GAXXXETAWACGXH 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	59.7	479	1 NAOX_MYCPN	P75389 mycoplasma
2	36	58.1	1533	1 PUM_DROME	P25822 drosophila
3	35	56.5	312	1 OACL_HUMAN	096kk4 homo sapien
4	35	56.5	662	1 DCBS_HUMAN	P19113 homo sapien
5	34	54.8	210	1 C2TA_HUMAN	P33076 homo sapien
6	34	54.8	269	1 PDX5_MOUSE	P99029 mus musculus
7	34	54.8	356	1 SURK_MOUSE	0643100 mus musculus
8	34	54.8	362	1 TRMA_NEIMA	P06579 bos taurus
9	34	54.8	404	1 TIRG_PIG	09t82 neisseria m
10	34	54.8	501	1 PHRL_SINAL	09n11 sus scrofa
11	34	54.8	640	1 Y051_MYCUDI	P40115 sinapis alba
12	34	54.8	672	1 HYFB_ECOLI	010880 mycobacteri
13	34	54.8	1107	1 ALA2_ARATH	P23492 escherichia
14	34	54.8	1612	1 TP2B_CRTIQ	P98205 arabidopsis
15	34	54.8	1626	1 TP2B_HUMAN	084399 cricetus
16	34	54.8	1627	1 TP2B_CHICK	Q04511 mus musculus
17	34	54.8	354	1 TFF1_ALCEU	Q02880 homo sapien
18	34	54.8	377	1 C1Q1L_XENIA	042131 gallus gallus
19	33	53.2	404	1 MOEA_HAEN	P27137 alcaligenes
20	33	53.2	404	1 KASI_STRNA	P02057 xenopus lae
21	33	53.2	421	1 SNAK_DRONE	P45210 haemophilus
22	33	53.2	430	1 MOTI_MOUSE	P16540 streptomyce
23	33	53.2	493	1 MOTI_CRTIQ	P05049 drosophila
24	33	53.2	494	1 MOTI_RAT	P53986 mus musculus
25	33	53.2	494	1 MOTI_HUMAN	P03064 cicatulus
26	33	53.2	494	1 RRPB_IBV	P53987 rattus norvegicus
27	33	53.2	500	1 MULI_THEFH	P53985 homo sapien
28	33	53.2	501	1 Y4QC_RHISN	P12723 avian infec
29	33	53.2	508	1 THIX YEAST	Q9ras4 thermus the
30	33	53.2	599	1 THIX YEAST	P55624 rhizobium s
31	33	53.2	599	1 CIQL_MOUSE	P08485 saccharomyces
32	33	53.2	604	1 CIQL_MOUSE	P08579 saccharomyces
33	33	53.2	604	1 P97414	P97414 mus musculus

ALIGNMENTS

RESULT 1

NAOX_MYCPN
ID NAOX_MYCPN
STANDARD: PRM; 479 AA.

P75389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable NADH oxidase (EC 1.6.99.3) (NOKase).
GN NOX OR MPN391 OR MP444.
OS Mycoplasma pneumoniae
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TAXID=2104;

[1]

SEQUENCE FROM N.A.
RL STRAIN=ATCC 29342 / MT29;
RC MEDLINE=97105885; PubMed=8948633;
RX Himmelreich R., Hilbert H., Plagens H., Pirk E., Li B.-C., Hermann R.;
RA "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae," Nucleic Acids Res. 24:4420-4449(1996).

CC -1- FUNCTION: CATALYZES THE FOUR-ELECTRON REDUCTION OF MOLECULAR OXYGEN TO WATER (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 NADH + O(2) = 2 NAD(+) + 2 H(2)O.
CC -1- COFACTOR: FAD.
CC -1- SIMILARITY: SOME, TO FLAVOPROTEIN OXIDOREDUCTASES. HIGHEST, TO NADH PEROXIDASE (NER).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
CC -----
DR HSSP; P37062; 1NHP;
DR InterPro; IPR001327; FAD_Pyr_redox.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF00070; Pyr_redox; 1.
DR ProDom; PD000139; FAD_Pyr_redox; 1.
DR ProDom; PD000139; FAD_Pyr_redox; 1.
DR Oxidoreductase; NAD; Flavoprotein; FAD; Redox-active center;
KW Complete proteome.
FT NP_BIND 3 33 FAD (ADP PART) (PROBABLE).
FT ACT_SITE 11 43 BASE (BY SIMILARITY).
FT ACT_SITE 43 43 REDOX-ACTIVE.
FT NP_BIND 170 185 NAD (BY SIMILARITY).
FT ND_BIND 295 305 FAD (FLAVIN PART) (BY SIMILARITY).
FT SEQUENCE 479 AA; 52875 MW; 6686EA3BA8E53F1 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 479;
Best local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

FT DOMAIN 259 271 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 272 291 7 (POTENTIAL).
 FT DOMAIN 292 312 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 96 188 BY SIMILARITY.
 FT CARBOHYD 4 4 N-LINKED (GLCNAC, . .) (POTENTIAL).
 SQ 312 AA; 34351 MW; B6FD600E700CB2CC CRR64;
 DR Query Match 56.5%; Score 35; DB 1; Length 312;
 DR Best Local Similarity 83.3%; Pred. No. 21;
 DR Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 7 TAWAGC 12
 Db 146 SAWAGG 151

RESULT 4
 DCHS_HUMAN STANDARD; PRT; 662 AA.

ID DCHS_HUMAN
 AC PI9113;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Histidine decarboxylase (EC 4.1.1.22) (HDC).
 GN HDC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OR NCBI_TAXID=9606;
 RN [1] SQUENCE FROM N.A.
 RK MEDLINE=91016941; PubMed=2216786;
 RA Yamauchi K., Rukiko S., Okikawa Y., Tanno Y., Maeyama K.,
 RA Watanabe T., Satoh K., Yoshizawa M., Shibahara S., Takishima T.;
 RK "Nucleotide sequence of the cDNA encoding L-histidine decarboxylase
 derived from human basophilic leukaemia cell line, KU-812-F.";
 RL Nucleic Acids Res. 18:5891-5891(1990).
 RN [2] SEQUENCE FROM N.A.
 RK MEDLINE=92119328; PubMed=1768863;
 RA Zahnow C.A., II H.F., McBride O.W., Joseph D.R.;
 RK "Cloning of the cDNA encoding human histidine decarboxylase from an
 erythroleukemia cell line and mapping of the gene locus to chromosome
 15." Seq. 1:395-400(1991).

RP SEQUENCE FROM N.A.
 RK MEDLINE=94117478; PubMed=8288622;
 RA Yatsunami K., Ohisu H., Tsuchikawa M., Higuchi T., Ishibashi K.,
 RA Shida A., Shima Y., Nakagawa S., Yamauchi K., Yamamoto M.,
 RA Hayashi N., Watanabe T., Ichikawa A.;
 RK "Structure of the L-histidine decarboxylase gene.";
 RL J. Biol. Chem. 269:1554-1559(1994).
 CC -I- CATALYTIC ACTIVITY: L-histidine - histamine + CO(2).
 CC -I- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -I- PATHWAY: Histamine biosynthesis.
 CC -I- SUBUNIT: HOMODIMER.
 CC -I- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYDC).
 CC -I- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; X54297; CAA38196.1; -;
 DR EMBL; M60445; AAC41698.1; -;
 DR EMBL; D16583; BAM04015.1; -;
 DR PIR; S11492; S11492.
 DR Genew; HGNC:4855; HDC.

RESULT 5
 C2PA_HUMAN STANDARD; PRT; 1130 AA.

ID C2PA_HUMAN
 AC P33076;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MHC class II transactivator CIITA.
 GN MHC2TA OR CIITA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1] NCBITAXID=9606;
 RK SEQUENCE FROM N.A., AND VARIANTS.
 RX MEDLINE=94006536; PubMed=8402895;
 RA Steimle V., Otten L.A., Zufferey M., Mach B.;
 RT "Complementation cloning of an MHC class II transactivator mutated in
 hereditary MHC class II deficiency (or bare lymphocyte syndrome)." ;
 RL Cell 75:135-146(1993).
 RN [2] SEQUENCE FROM N.A.
 RK MEDLINE=95265136; PubMed=7749984;
 RA Riley J.L., Westerheide S.D., Price J.A., Brown J.A., Boss J.M.;
 RT "Activation of class II MHC genes requires both the κ box region and
 the class II transactivator (CIITA)." ;
 RL Immunity 2:533-543(1995).
 CC -I- FUNCTION: ESSENTIAL FOR TRANSCRIPTIONAL ACTIVITY OF THE HLA CLASS
 II PROMOTER: ACTIVATION IS VIA THE PROXIMAL PROMOTER. NO DNA
 BINDING OF IN VITRO TRANSLATED CIITA WAS DETECTED. MAY ACT IN A
 CONTRACTIVATOR-LIKE FASHION THROUGH PROTEIN-PROTEIN INTERACTIONS BY
 CONTACTING FACTORS BINDING TO THE PROXIMAL MHC CLASS II PROMOTER,
 TO ELEMENTS OF THE TRANSCRIPTION MACHINERY, OR BOTH. ALTERNATIVELY,
 IT MAY ACTIVATE HLA CLASS II TRANSCRIPTION BY MODIFYING PROTEINS
 THAT BIND TO THE MHC CLASS II PROMOTER.
 CC -I- DISEASE: DEFECTS IN MHC2TA ARE A CAUSE OF HEREDITARY MHC CLASS II
 DEFICIENCY (ALSO KNOWN AS BARE LYMPHOCYTE SYNDROME (BLS) OR HLA
 CLASS II-DEFICIENT COMBINED IMMUNODEFICIENCY); A FORM OF SEVERE
 CC COMBINED IMMUNODEFICIENCY DISEASE (SCID). MHC2TA IS LINKED WITH
 CC BLS COMPLEMENTATION GROUP A.
 CC -I- SIMILARITY: CONTAINS 1 NACHT DOMAIN.
 CC -I- DATABASE NAME=Atlas Genet. Cytogenet. Oncol. Haematol.:
 CC WWW="<http://www.intobioigen.fr/services/chromcancer/Genes/MHC2TAID260.html>".
 CC -I-
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; U18259; AAA88861.1; -;

DR Genew; HNC:7067; MHC2TA.
 DR MIM: 600005; -.
 DR 209920; -.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003592; LRR_out.
 DR Pfam; PF00560; LRR; 1.
 DR SMART; SM00370; LRR; 4.
 DR PROSITE; PS50837; NACHT; 1.
 DR transcription regulation; Activator; Arp-binding; Disease mutation;
 SCID.
 KW DOMAIN 414 724 NACHT.
 FT DOMAIN 52 137 ASP/GLU-RICH (ACIDIC).
 NP_BIND 420 427 ATP (POTENTIAL).
 FT VARIANT 120 120 K-> IE (IN BLS).
 FT VARIANT 500 500 A-> G (IN BLS).
 FT VARIANT 940 963 MISSING (IN BLS).
 FT /FTID=VAR_005129.
 SQ SEQUENCE 1130 AA; 123456 MW; DC4D0B1802987E05 CRC64;
 Query Match Score 35; DB 1; Length 1130;
 Best Local Similarity 56.5%; Pred. No. 66; Mismatches 0; Indels 5; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 AXXXTETAWACG 12
 Db 430 AGAVGSRRAWACG 440
 DR [6]
 RESULT 6
 PDX5_MOUSE STANDARD; PRT; 210 AA.
 ID PDX5_MOUSE STANDARD; PRT; 210 AA.
 AC P99029; Q9QZ75; Q9QX45; [1]
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Peroxiredoxin 5, mitochondrial precursor (Ppx-V) (Peroxisomal
 antioxidant enzyme) (PLP) (Thioredoxin peroxidase PMP20) (Antioxidant
 enzyme B166) (AOEB166) (Liver tissue 2D-page spot 2D-0014IV).
 DE PRDx6 OR PRDx6
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=20145535; Pubmed=10679306;
 RA Zhou Y., Kok K.H., Chun A.C.S., Wong C.M., Wu H.W., Lin M.C.M.,
 RA Fung P.C., Kung H.-F., Jin D.-Y.,
 RT "Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits
 P53-induced apoptosis.";
 RL Biochem. Biophys. Res. Commun. 268:921-927(2000).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=99445545; Pubmed=1054471;
 RA Yamashita H., Avraham S., Jiang S., London R., Van Veldhoven P.P.,
 RA Subramani S., Rogers R.A., Avraham H.;
 RT "Characterization of human and murine PMP20 peroxisomal proteins that
 exhibit antioxidant activity in vitro.";
 RL J. Biol. Chem. 274:29897-29904(1999).
 RN [3]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=CH/HeJ; TISSUE=Lung;
 RX MEDLINE=9945229; Pubmed=10531424;
 RA Krooops B., Cliffe A., Bogard C., Arsalane K., Wattiez R., Hermans C.,
 RA Duconselle E., Faligagne P., Bernard A.,
 RT "Cloning and characterization of AOEB166, a novel mammalian
 antioxidant enzyme of the peroxiredoxin family.",
 J. Biol. Chem. 274:30451-30458(1999).
 RX SEQUENCE FROM N.A.
 RX MEDLINE=20218665; Pubmed=10733630;

RA Lee T.H., Kim S.J., Kang S.W., Lee K.K., Rhee S.G., Yu D.Y.;
 RT "Molecular cloning and characterization of the mouse Peroxiredoxin V
 gene";
 RL Biochem. Biophys. Res. Commun. 270:356-362(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawaji J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Harada A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alzava K., Izawa M., Nishii K., Kyosawa H., Kondo S., Yamamoto I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batyalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matuo Y., Nikaiko I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarulli J., Mombaart P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitzker C., Whittaker C., Wilming L.,
 RA Wysniaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.; Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE OF 50-61.
 RC TISSUE=Liver;
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G.J., Yan J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Cowhorne M.;
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: Reduces hydrogen peroxide and alkyl hydroperoxides with
 CC reducing equivalents provided through the thioredoxin system.
 CC -!- INVOLVED: In intracellular redox signaling.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial, peroxisomal and cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS: A SINGLE NUCLEAR GENE PRODUCES BOTH FORMS
 CC BY USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 7.6. ITS MW IS: 14.6 kDa.
 CC -!- SIMILARITY: BELONGS TO THE PEROXIREDOXIN 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL_AF197951; AR04855; 1; -.
 DR EMBL_AF124994; AR02732; 1; -.
 DR EMBL_AF11033; AR01350; 1; -.
 DR EMBL_AF208730; AR01106; 1; JOINED.
 DR EMBL_AF208729; AR021016; 1; JOINED.
 DR EMBL_AK002363; BAB22058; 1; -.
 DR EMBL_AK00332; BAB22720; 1; -.
 DR EMBL_BC008174; AR008174; 1; -.
 DR HSSP; 3004; 1HD2
 DR MGD; MG1.1859821; Prdx6.
 DR SWISS-2DPAGE; PR99029; MOUSE.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR Pram; PF00578; Ahpc-TSA; 1.

DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS01187; EGF_CA; 2.
 KW Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;
 Glycoprotein; EGF-like domain.
 FT NON-TER 1 1
 FT DOMAIN 1 296
 FT TRANSMEM 297 320
 FT DOMAIN 321 356
 FT DOMAIN 17 57
 FT DOMAIN 60 98
 FT DOMAIN 99 137
 FT DOMAIN 139 179
 FT DOMAIN 178 214
 FT DOMAIN 215 254
 FT DISULFID 21 32
 FT DISULFID 28 41
 FT DISULFID 43 56
 FT DISULFID 64 72
 FT DISULFID 68 82
 FT DISULFID 84 97
 FT DISULFID 103 114
 FT DISULFID 110 123
 FT DISULFID 125 136
 FT DISULFID 143 152
 FT DISULFID 148 162
 FT DISULFID 164 178
 FT DISULFID 182 191
 FT DISULFID 187 199
 FT DISULFID 201 213
 FT DISULFID 219 228
 FT DISULFID 224 237
 FT DISULFID 239 253
 FT CARBOHYD 271
 SQ SEQUENCE 356 AA; 37795 MW; 29841F057ABE093 CRC64;
 Best Local Similarity 54.8%; Score 34; DB 1; Length 356;
 Matches 5; Conservative 100%; Mismatches 0; Indels 0; Gaps 0;
 Qy 8 AWACG 12
 Db 18 AWACG 22

RESULT 9

ID	TRMA_NEIMA	STANDARD;	PRT;	362 AA.
Q9J82;				15-JUN-2002 (Rel. 41, created)
DT				15-JUN-2002 (Rel. 41, last sequence update)
DE	trna (uracil-5'-methyltransferase (EC 2.1.1.35) (tRNA(M-5-U54) - methyltransferase) (RUMY). GN OR tRNA58.			15-JUN-2002 (Rel. 41, last annotation update)
NEISSERIA MENINGITIS	Neisseria meningitidis (serogroup A).			15-JUN-2002 (Rel. 41, last annotation update)
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			15-JUN-2002 (Rel. 41, last annotation update)
OX	NCBI_TaxID=65699;			15-JUN-2002 (Rel. 41, last annotation update)
RN [1]	SEQUENCE FROM N. A.			15-JUN-2002 (Rel. 41, last annotation update)
RC	STRAIN=22491 / Serogroup A / Serotype 4A;			15-JUN-2002 (Rel. 41, last annotation update)
RX	MEDLINE=0022556; PubMed=10761919;			15-JUN-2002 (Rel. 41, last annotation update)
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davis R.M., Devlin K., Feltwell T., Hamlin N., Holroyd S., Jagels K., Leather S., Moulé S., Mungall K., Quail M.A., Ratnandran M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G.; Barrell B.G.;			15-JUN-2002 (Rel. 41, last annotation update)
RT	"Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491";			15-JUN-2002 (Rel. 41, last annotation update)
RL	Nature 404:502-506 (2000).			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position 54 (M-5-U54) in all tRNA (By similarity).			15-JUN-2002 (Rel. 41, last annotation update)

RESULT 10

ID	I1BC_PIG	STANDARD;	PRT;	404 AA.
AC	Q9N211;			15-JUN-2002 (Rel. 41, created)
DT				15-JUN-2002 (Rel. 41, last sequence update)
DE	Interleukin-1 beta converting enzyme precursor (IL-1BC) (EC 3.4.22.36) (IL-1 beta converting enzyme) (ICE) (Interleukin-1 beta converting enzyme (p45)) (Caspase-1) (CASP-1).			15-JUN-2002 (Rel. 41, last annotation update)
DR	Caspase-1 (CASP-1) (p45).			15-JUN-2002 (Rel. 41, last annotation update)
DE	DR GN CASP1 OR IL1BC.			15-JUN-2002 (Rel. 41, last annotation update)
OS	Sus scrofa (Pig).			15-JUN-2002 (Rel. 41, last annotation update)
OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			15-JUN-2002 (Rel. 41, last annotation update)
OX	OX NCBL_TaxID=9823;			15-JUN-2002 (Rel. 41, last annotation update)
RN [1]	SEQUENCE FROM N. A.			15-JUN-2002 (Rel. 41, last annotation update)
RP	SEQUENCE FROM N. A.			15-JUN-2002 (Rel. 41, last annotation update)
RX	MEDLINE=20039779; PubMed=10574622;			15-JUN-2002 (Rel. 41, last annotation update)
RA	Muneta Y., Shinohji Y., Yokomizo Y., Mori Y.; "Molecular cloning of porcine interleukin-1beta converting enzyme and differential gene expression of IL-1beta converting enzyme in porcine alveolar macrophages.";			15-JUN-2002 (Rel. 41, last annotation update)
RT	RJ Interferon Cytokine Res. 19:1289-1296(1999).			15-JUN-2002 (Rel. 41, last annotation update)
RL	-!- FUNCTION: Thiol protease that cleaves IL-1 beta between an ASP and AN ALA, releasing the mature cytokine which is involved in a variety of inflammatory processes (By similarity).			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- CATALYTIC ACTIVITY: Release of interleukin 1-beta by specific cleavage at 116-Asp-1-Ala-117 and 27-Asp-1-Gly-28 bonds in precursor. Also hydrolyzes the small-molecule substrate, Ac-Tyr-Val-Ala-Asp-1-NHMe.			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- SUBUNIT: Tetramer that consists of two heterodimers of a 20 kDa (P20) and a 10 kDa (P10) subunits. P20 can also form a heterodimer with the epsilon isoform which then has an inhibitory effect (By similarity).			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- PTM: The two subunits are derived from the precursor sequence by an autocatalytic mechanism.			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- SIMILARITY: BELONGS TO PEPTIDE FAMILY C14.			15-JUN-2002 (Rel. 41, last annotation update)
CC	-!- SIMILARITY: CONTAINS 1 CARD DOMAIN.			15-JUN-2002 (Rel. 41, last annotation update)

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AB027296; BAA9531.1;
 DR HSSP: P29466; ICE.
 DR MEROPS: C14.001;--.
 DR InterPro: IPR001315; CARD.
 DR InterPro: IPR002398; ICE.
 DR InterPro: IPR002138; ICE_P10.
 DR InterPro: IPR001309; ICE_P20.
 DR Pfam: PF00519; CARD; 1.
 DR Pfam: PF00555; ICE_P10; 1.
 DR Pfam: PF0056; ICE_P20; 1.
 DR PRINTS; PRO0376; ILBRCENZYME.
 DR SMART; SM00114; CARD; 1.
 DR PROSITE; PS50209; CARD; 1.
 DR PROSITE; PS01124; CASPASE_CYS; 1.
 DR PROSITE; PS01211; CASPASE_HIS; 1.
 DR PROSITE; PS50207; CASPASE_P10; 1.
 DR PROSITE; PS50208; CASPASE_P20; 1.
 KW Hydrolase; Thiol protease; Zymogen.
 FT CHAIN 120 297 POTENTIAL
 FT PROPEP 1 119 INTERLEUKIN-1 BETA CONVERTASE P10.
 FT DOMAIN 1 91 CARD.
 FT ACT_SITE 237 237 BY SIMILARITY.
 FT ACT_SITE 285 285 BY SIMILARITY.
 SQ SEQUENCE 404 AA; 44881 MW; 437DC787E85FB449 CRC64;

Query Match 54.8%; Score 34; DB 1; Length 404;
 Best Local Similarity 71.4%; Pred. No. 40; Mismatches 1; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Misnmatches 1; Indels 0; Gaps 0;

QY 6 ETAWACG 12

Db 359 EHAWSCG 365

RESULT 11

ID	PHRL_SINAL	STANDARD:	PRT:	501 AA.
AC	P4015;			31. Created)
DT	01-FEB-1995	(Rel.	31,	Last sequence update)
DT	15-DEC-1998	(Rel.	37,	Last annotation update)
DE	Deoxyribonucleic acid photolyase (EC 4.1.99.3)	(DNA photolyase)		
DE	(Photoreactivating enzyme).			
GN	PHRL.			
OS	sinapis alba (white mustard) (Brassica hirta).			
OC	Sinapycrota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosidia II; Brassicales; Brassicaceae; Sinapis.			
OC	NCBI TaxID=3728;			
[1]	SEQUENCE FROM N.A.			
RK	MEDLINE-#94073224; PubMed=8252071;			
RA	Batschauer A.;			
RT	"A plant gene for photolysse: an enzyme catalyzing the repair of UV-			
RT	light-induced DNA damage.";			
RU	Plant J. 4:705-709 (1993).			
CC	- FUNCTION: THIS ENZYME CATALYZES THE LIGHT-DEPENDENT MONOMERIZATION			
CC	(300-600 NM) OF CYCLOBUTYL PYRIMIDINE DIMERS (IN CIS-SYN			
CC	CONFIGURATIONS), WHICH ARE FORMED BETWEEN ADJACENT BASES ON THE			
CC	SAME DNA STRAND, UPON EXPOSURE TO ULTRAVIOLET RADIATION.			
CC	- CATALYTIC ACTIVITY: CYCLOBUTADIPYRIMIDINE (IN DNA) = 2 pyrimidine			
CC	residues (in DNA).			

CC -1 COFACTOR: CONTAINS 2 CHROMOPHORES: A REDUCED FLAVIN (FADH2) AND AN
 CC OXIDIZED 8-HYDROXY-5,6-DIHYDROFLAVIN (F420). BOTH CHROMOPHORES ARE
 CC BOUND BY NON-COVALENT INTERACTIONS (PROBABLE).
 CC
 CC -1 INDUCTION: BY VISIBLE LIGHT.
 CC
 CC -1 SIMILARITY: BELONGS TO THE DNA PHOTOLYASE CLASS-1 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X72019; CA50898; 1;
 DR InterPro: IPR000474; DNA_photolyase_1.
 DR InterPro: IPR002081; DNA_photolyase_1.
 DR InterPro: IPR005101; FAD_binding_7.
 DR Pfam: PF00815; DNA_photolyase; 1.
 DR Pfam: PF03411; FAD_binding_7.
 DR Prodrom: PD004390; DNA_photolyase; 1.
 DR PROSITE: PS00394; DNA_PHOTOLYASES_1; 1.
 DR PROSITE: PS00691; DNA_PHOTOLYASES_1; 1.
 KW Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding;
 SQ Sequence: CHROMOPHORE; Flavoprotein; FAD; DNA repair; DNA-binding;
 SQ Y051_MVCTU 501 AA; 57021 MW; C6B25CE0A33B90DB CRC64;
 Query Match 54.8%; Score 34; DB 1; Length 501;
 Best Local Similarity 83.3%; Pred. No. 48; Mismatches 1; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Misnmatches 1; Indels 0; Gaps 0;

QY 6 ETAWAC 11

Db 181 ETWAC 186

RESULT 12

ID	Y051_MVCTU	STANDARD:	PRT:	640 AA.
AC	Y051_MVCTU			
DT	01-0880; 053628;			
DT	01-OCT-1996 (Rel. 34; Created)			
DT	15-DEC-1998 (Rel. 37; Last sequence update)			
DT	15-JUN-2002 (Rel. 41; Last annotation update)			
DE	Hypothetical protein Rv0083.			
GN	RV0083 OR MT0090 OR MVCTY251.01 OR MTW030.27.			
OC	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI TaxID=1773;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPTRAIN-H37RV;			
RX	MEDLINE-#9825987; Pubmed=9634230;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglemeier K., Gas J., Barry C.E. III, Tekla F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutten S., Seeger K., Skelton S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;			
RA	"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";			
RT	Nature 393:537-544 (1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPTRAIN-CDC 1551 / Oshkosh;			
RA	Fleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O., Peierson J., Deboy R., Dodson R., Gwinn M.L., Hafft D., Hickley E., Kolonay J.F., Nelson W.C., Umayam L.A., Embley M.D., Salzberg S.L., DeJcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;			
RA	"Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";			

RL	Science 277:1453-1474(1997).
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC	-!- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC	OF CHLOROPLASTS OR MITOCHONDRIA.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; A1021428; CAA16264_1; .
DR	EMBL; AE006920; ARK44315_1; .
DR	TIGR; MT0090; .
DR	Tuberculast; Rv0083; .
DR	InterPro; IPR001750; Oxidored_q1.
DR	Pfam; PF00361; oxidored_q1; 1.
KW	Hypothetical protein; Oxidoreductase; Transmembrane;
KW	Complete proteome.
FT	TRANSMEM 8 POTENTIAL.
FT	TRANSMEM 28 POTENTIAL.
FT	TRANSMEM 52 POTENTIAL.
FT	TRANSMEM 90 POTENTIAL.
FT	TRANSMEM 136 POTENTIAL.
FT	TRANSMEM 156 POTENTIAL.
FT	TRANSMEM 179 POTENTIAL.
FT	TRANSMEM 208 POTENTIAL.
FT	TRANSMEM 228 POTENTIAL.
FT	TRANSMEM 241 POTENTIAL.
FT	TRANSMEM 277 POTENTIAL.
FT	TRANSMEM 297 POTENTIAL.
FT	TRANSMEM 318 POTENTIAL.
FT	TRANSMEM 352 POTENTIAL.
FT	TRANSMEM 372 POTENTIAL.
FT	TRANSMEM 391 POTENTIAL.
FT	TRANSMEM 411 POTENTIAL.
FT	TRANSMEM 446 POTENTIAL.
FT	TRANSMEM 466 POTENTIAL.
FT	TRANSMEM 497 POTENTIAL.
FT	TRANSMEM 517 POTENTIAL.
FT	TRANSMEM 619 POTENTIAL.
SQ	SEQUENCE 640 AA; 65627 MW; 6254C299DBD108A8 CRC64;
Query Match	54.8%; Score 34; DB 1; Length 640;
Best local Similarity	83.3%; Pred. No. 60;
Matches	5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	9 WACGXA 14
Db	533 WACGCA 538
RESULT 13	
HYPER_ECOLI	STANDARD; PRT; 672 AA.
ID	HYPER_ECOLI
AC	P23482;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Hydrogenase-4 component B (EC 1.-.-.-).
GN	HFB_OB2402.
OS	Escherichia coli.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC	Escherichia.
NCBI_TAXID	562;
[1]	
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC	STRAIN=K12;
RA	Andrews S.C., McCay J., Ambler A., O'Neill M., Berks B.C., Guest J.R.,
RL	Submitted (Oct-1996) to the EMBL/GenBank/DDBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=K12 / MG1655;
RC	MEDLINE:97426617; PubMed=9270503;
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J.J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA	Mau B., Shao Y.;
RT	"The complete genome sequence of Escherichia coli K-12."
RL	Science 277:1453-1474(1997).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=K12;
RX	Medline=97349980; PubMed=9205837;
RA	Itoh T., Kimura S., Kiragawa M., Makino K., Inada T., Isono K., Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H., Oshima T., Okuma S., Saito N., Sampei G., Satoh Y., Sivasundaram S., Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C., Yamagata S., Horikuchi T.,
RA	"Construction of a contiguous 874-kb sequence of the Escherichia coli K12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features.";
RL	DNA Res. 4:91-113(1997).
RN	[4]
RP	SEQUENCE OF 1-34 FROM N.A.
RC	STRAIN=K12;
RX	Medline=91202105; PubMed=2016588;
RA	Andrews S.C., Harrison P.M., Guest J.R.; "A molecular analysis of the 53.3 minute region of the Escherichia coli linkage map";
RT	J. Gen. Microbiol. 137:361-367(1991).
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (By similarity).
CC	-!- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC	OF CHLOROPLASTS OR MITOCHONDRIA.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M63654; AAB8564_1; .
DR	EMBL; AE00335; AAC75525_1; .
DR	EMBL; D90876; BAA16360_1; .
DR	EMBL; D90877; BAA16370_1; .
PIR	PIR; D49749; D49749.
DR	EcoGene; EG11282; hyfb.
DR	InterPro; IPR001750; Oxidored_q1.
DR	Pfam; PF00361; oxidored_q1; 1.
KW	Oxidoreductase; Transmembrane; Inner membrane; Complete proteome.
FT	TRANSMEM 6 POTENTIAL.
FT	TRANSMEM 26 POTENTIAL.
FT	TRANSMEM 31 POTENTIAL.
FT	TRANSMEM 80 POTENTIAL.
FT	TRANSMEM 120 POTENTIAL.
FT	TRANSMEM 140 POTENTIAL.
FT	TRANSMEM 165 POTENTIAL.
FT	TRANSMEM 200 POTENTIAL.
FT	TRANSMEM 243 POTENTIAL.
FT	TRANSMEM 273 POTENTIAL.
FT	TRANSMEM 312 POTENTIAL.
FT	TRANSMEM 343 POTENTIAL.
FT	TRANSMEM 363 POTENTIAL.
FT	TRANSMEM 385 POTENTIAL.
FT	TRANSMEM 405 POTENTIAL.
FT	TRANSMEM 428 POTENTIAL.
FT	TRANSMEM 448 POTENTIAL.
FT	TRANSMEM 475 POTENTIAL.
FT	TRANSMEM 495 POTENTIAL.
FT	TRANSMEM 505 POTENTIAL.
FT	TRANSMEM 525 POTENTIAL.
FT	TRANSMEM 532 POTENTIAL.
FT	TRANSMEM 652 POTENTIAL.
SQ	SEQUENCE 672 AA; 72582 MW; 3CA8C2BD25E54B1B CRC64;
Query Match	54.8%; Score 34; DB 1; Length 672;
Best local Similarity	100.0%; Pred. No. 63;
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	8 AWACG 12
Db	565 AWACG 569
RESULT 14	

KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
FT NP_BIND 170 ATP (POTENTIAL).
FT ACT_SITE 814 DNA CLEAVAGE (BY SIMILARITY).
SEQUENCE 1612 AA; 182074 MW; C01D6RC40620FC6B CRC64;
SQ

Query Match 54.8%; Score 34; DB 1; Length 1612;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GAXXXETAWAC 11
|||
Db 861 GAEGIGTGWAC 871

Search completed: March 25, 2003, 08:20:12
Job time : 10.4848 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

ON protein - protein search, using sw model

Run on: March 25, 2003, 08:19:47 ; Search time 30.9697 seconds
(without alignments)
93.145 Million cell updates/sec

Title: US-09-646-532B-2

Perfect score: 62

Sequence: 1 GAXXETAWACGXA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 20604715 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

SPTRMBL 21: *
1: sp_archea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodont: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriapl: *
17: sp_archeapl: *

Pred. No. 1 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	46	74.2	290	16 Q8UIQ4	Q8uiq4 agrobacteri
2	40	64.5	276	5 Q9J398	Q9J398 phaedon coc
3	40	64.5	423	2 Q9KH6	Q9kh6 streptomyce
4	39	62.9	719	3 Q8WJ0	Q8wj0 pneumocysti
5	39	62.9	1681	4 Q9GJ1	Q9gj1 homo sapien
6	38	61.3	195	16 Q92K85	Q92k85 rhizobium m
7	38	61.3	229	9 Q9XV2	Q9xv2 streptococc
8	38	61.3	229	9 Q9KH2	Q9kh2 streptococc
9	38	61.3	741	16 Q9HZM0	Q9hzm0 pseudomonas
10	37	59.7	218	16 Q8A5N1	Q8a5n1 caulobacter
11	37	59.7	275	16 Q9RXV3	Q9rxv3 delinococcus
12	37	59.7	334	16 Q9KAP4	Q9kap4 streptomyce
13	37	59.7	448	2 Q9RH01	Q9rh01 lactobacill
14	37	59.7	502	16 Q8Ab47	Q8ab47 escherichia
15	37	59.7	670	2 Q8NNNT	Q8nnnt escherichia
16	59.7	707	16 Q9EYD1	"genome sequence of the plant pathogen and biotechnology agent Q9eydl escherichia	

RESULT 1	ID	Q8UIQ4	PRELIMINARY;	PRT;	290 AA.
	Q8UIQ4				
	AC	Q8UIQ4;			
	DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
	DE	transcriptional regulator, AraC family.			
	GN	ATU0239 OR AGR_C-407.			
	OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).			
	OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
	OC	Rhizobiaceae; Rhizobium.			
	OX	NCBI_TaxID=176299;			
	RN	[1]			
	RP	SEQUENCE FROM N.A.			
	RX	MEDLINE=212108550; PubMed=11743193;			
	RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Almeida N.F. JR., Woo L., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. JR., Bovee D. SR., Chen Y., Paules I.T., Eisen J.A., Karp P.D., Bovee D. SR., Chapman P., Clendening J., Deatherage G., Gillett W., Grant C., Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri R., Raymond C., Rouse G., Saengphimmacak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Bidle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W., RT			
	RA	RT "the genome of the natural genetic engineer Agrobacterium tumefaciens C58," Science 294:2317-2323(2001).			
	RL	SEQUENCE FROM N.A.			
	RN	MEDLINE=21208551; PubMed=11743194;			
	RA	Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Qurollo B., Goldhan B.S., Gao Y., Askren M., Halling C., Mullin L., Homicel K., Gordom J., Vaudin M., Tartchouk O., Epp A., Liu F., Wolam C., Allinger M., Dougherty D., Scott C., Lappas C., Marfuzz B., Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G., Cleo C., Slatter S.,			
	RT	"genome sequence of the plant pathogen and biotechnology agent Q9eydl escherichia			

RT	Agrobacterium tumefaciens C58. ";	OC	Actinomycetales; Streptomyceae; Streptomyces;
RL	Science 299:2323-2328(2001).	OX	NCBI_TaxID=115628;
DR	EMBL; AED08996; AAU41261.1; -.	RN	[1]
DR	EMBL; AE007963; AAK86055.1; ALT_INIT.	RP	SEQUENCE FROM N.A.
KW	Complete Proteome.	RA	Piel J.; Hoang K.; Moore B.S.;
SQ	SEQUENCE 290 AA; 32633 MW; CEF3643PA611B58A CRC64;	RT	"Natural Metabolic Diversity Encoded by the Enterocin Biosynthesis
Query Match	74.2%; Score 46; DB 16; Length 290;	RT	Gene Cluster";
Best Local Similarity	64.3%; Pred. No. 1.5;	RL	J. Am. Chem. Soc. 122:5415-5416(2000).
Matches	9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	DR	EMBL; AF254925; AAU81728.1; -.
Qy	1 GAXXETAWACGX 14	DR	HSSP; P7283; JE5M.
Db	GASLAETAFACGFA 264	DR	InterPro; IPR0094; Ketocacyl-synt.
RESULT 2		DR	Pfam; PF00109; ketocacyl-synt; 1.
097398	PRELIMINARY; PRT; 276 AA.	DR	Pfam; PF02801; ketocacyl-synt_C; 1.
ID	097398	DR	PROSITE; PS00606; _KETOACYL_SYNTHASE; UNKNOWN_1.
AC	097398;	SQ	SEQUENCE 423 AA; 44640 MW; A98CC6B2501F7DA CRC64;
DT	01-MAY-1999 (TREMBREL. 10, Created)	Qy	1 GAXXETAWACG 12
DT	01-MAY-1999 (TREMBREL. 10, Last sequence update)	Db	147 GSLARETAWAG 158
DT	01-JUN-2002 (TREMBREL. 21, Last annotation update)		
DE	Chymotrypsin precursor.		
OS	Phaedon cochleariae (Mustard beetle).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;		
OC	NCBI_TaxID=80249;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC			
RA	Gillard C., Jouanin L.; Submitted (Aug-1998) to the EMBL/GenBank/DDBJ databases. ALSO KNOWN AS THE TRYPSIN FAMILY.		
RA	TRYPSIN, Y17904; CAA6928.1; -.		
DR	HSSP; P00761; LEPT.		
DR	MEROPS; S01.122; -.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	PRINTS; PR0089; TRYPSIN; 1.		
DR	SMART; SNN0020; TRYSP_SPC; 1.		
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00135; TRYPSIN_HIS; UNKNOWN_1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
KW	Hydrolase; Serine protease; Signal.		
FT	SIGNAL 1		
FT	CHAIN 16		
FT	POTENTIAL.		
SEQUENCE	276 AA; 29868 MW; FC5FD05DB882A1DE CRC64;		
Qy	64.5%; Score 40; DB 5; length 276; Best Local Similarity 54.5%; Pred. No. 17; Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;		
Qy	2 AXXTANACG 12		
Db	65 ASAGETSWFG 75		
RESULT 3			
Q9HKG6	PRELIMINARY; PRT; 423 AA.		
ID	Q9HKG6		
AC	Q9HKG6; 01-OCT-2000 (TREMBREL. 15, Created)		
AC	01-OCT-2000 (TREMBREL. 15, Last sequence update)		
DT	01-DEC-2001 (TREMBREL. 19, Last annotation update)		
DE	Putative keto synthase alpha EnCA.		
GN	ENCA.		
OS	Streptomyces maritimus.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
RESULT 4			
Q96WJ0	PRELIMINARY; PRT; 719 AA.		
ID	Q96WJ0		
AC	Q96WJ0; 01-DEC-2001 (TREMBREL. 19, Created)		
DT	01-DEC-2001 (TREMBREL. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBREL. 21, Last annotation update)		
DE	Lanosterol synthase.		
GN	ERG7.		
OS	Pneumocystis carinii.		
OC	Eukaryote; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;		
OC	Pneumocystis.		
OX	NCBI_TaxID=4754;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Joubert B.M., Matsuda S.P.T.;		
RT	"Cloning and characterization of an oxisosqualene cyclase cDNA established lanosterol as a pneumocystis carinii metabolite.",		
RT	Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.		
RT	EMBL; AF28525; AAK8293.1; -.		
DR	InterPro; IPR001330; Prenyltrans.		
DR	Pfam; PF00432; prenyltrans; 3.		
SQ	SEQUENCE 719 AA; 83508 MW; 4904DA1B8829F675 CRC64;		
Qy	62.9%; Score 39; DB 3; length 719; Query Match		
Best Local Similarity	54.5%; Pred. No. 68; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;		
Qy	1 GAXXETAWAC 11		
Db	:1111 639 GSQVWQTAWC 649		
RESULT 5			
Q9UGJ1	PRELIMINARY; PRT; 1681 AA.		
ID	Q9UGJ1		
AC	Q9UGJ1; 01-MAY-2000 (TREMBREL. 13, Created)		
DT	01-MAY-2000 (TREMBREL. 13, Last sequence update)		
DT	01-JUN-2002 (TREMBREL. 21, Last annotation update)		
DE	RB-binding protein.		
GN	RBB2HIA.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OC	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

RA Kashuba V.; Protopopov A.; Podowski R.; Zabavovsky E.;
 RT "Isolation and chromosomal localization of a new human retinoblastoma
 binding protein 2 homolog 1a (RBBP2H1A).";
 RI Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AU243706; CAB63108.; -.
 DR InterPro; IPR001606; ARID.
 DR InterPro; IPR003347; TF_JM1C.
 DR InterPro; IPR003449; TF_JM1N.
 DR InterPro; IPR00198; Znf_C5HC2.
 DR InterPro; IPR001965; Znf_PRD.
 DR Pfam; PF01388; ARID; 1.
 DR Pfam; PF02373; JmjN; 1.
 DR Pfam; PF00628; PHD; 3.
 DR Pfam; PF02288; zf-C5HC2; 1.
 DR SMART; SM00501; BRIGHT; 1.
 DR SMART; SM00545; JmjN; 1.
 DR SMART; SM00249; PHD; 3.
 SQ SEQUENCE 1681 AA; 190118 MW; 0429FA1E7B4FB702 CRC64;
 Query Match 62.9%; Score 39; DB 4; Length 1681;
 Best Local Similarity 100.0%; Pred. No. 1.6e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 TAWAGC 12
 Db 77 TAWAGG 82

RESULT 6

Q92K85 PRELIMINARY; PRT; 195 AA.

ID Q92K85
 AC Q92K85
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical transmembrane protein Smc04211.
 GN R01881 or SMC04211.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID:382;
 RN [1]
 RB SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Aupre F., Batut J.,
 RA Bolard P., Becker M., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire E., Dreano S.,
 RA Pohl T., Portetelle D., Pueler A., Purnelle B., Ramspeger U.,
 RA Renard C., Thebaud P., Vandenberg M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 sinorhizobium meliloti strain 1021.";
 RU Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591788; CAC64640; 1; -.
 DR InterPro; IPR000080; SNase.
 DR Pfam; PF00565; SNase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 195 AA; 21051 MW; 6B7635673D5031FF CRC64;

Query Match 61.3%; Score 38; DB 16; Length 195;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 TAWAGC 12
 Db 69 TAWDCG 75

RESULT 7

Q9XJV2 PRELIMINARY; PRT; 229 AA.

AC Q9XJV2
 DR InterPro; IPR001965; Znf_PRD.
 DR InterPro; IPR003347; TF_JM1C.
 DR InterPro; IPR003449; TF_JM1N.
 DR InterPro; IPR00198; Znf_C5HC2.
 DR InterPro; IPR001965; Znf_PRD.

RI Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR InterPro; IPR001606; ARID.
 DR InterPro; IPR003347; TF_JM1C.
 DR InterPro; IPR003449; TF_JM1N.
 DR InterPro; IPR00198; Znf_C5HC2.
 DR InterPro; IPR001965; Znf_PRD.

DR Pfam; PF01388; ARID; 1.
 DR Pfam; PF02373; JmjN; 1.
 DR Pfam; PF00628; PHD; 3.
 DR Pfam; PF02288; zf-C5HC2; 1.
 DR SMART; SM00501; BRIGHT; 1.
 DR SMART; SM00545; JmjN; 1.
 DR SMART; SM00249; PHD; 3.
 SQ SEQUENCE 1681 AA; 190118 MW; 0429FA1E7B4FB702 CRC64;
 Query Match 62.9%; Score 39; DB 4; Length 1681;
 Best Local Similarity 100.0%; Pred. No. 1.6e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 TAWAGC 12
 Db 77 TAWAGG 82

RESULT 8

Q9MCCH2 PRELIMINARY; PRT; 229 AA.

ID Q9MCCH2
 AC Q9MCCH2
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE GP229;
 OS Streptococcus thermophilus bacteriophage SF18.
 OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID:74382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99412383; PubMed=10482618;
 RA Lucchini S., Desiere F., Brussow H.;
 RT "Comparative genomics of Streptococcus thermophilus phage species
 supports a modular evolution theory.";
 RL J. Virol. 73:647-865(1999).
 DR EMBL; AF158601; AAF63082.1; -.

SQ SEQUENCE 229 AA; 25682 MW; ED2AP9129827BD95 CRC64;
 Query Match 61.3%; Score 38; DB 9; Length 229;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 GAXXXETAWACXA 14
 Db 118 GARVYGNAWVCGNA 131

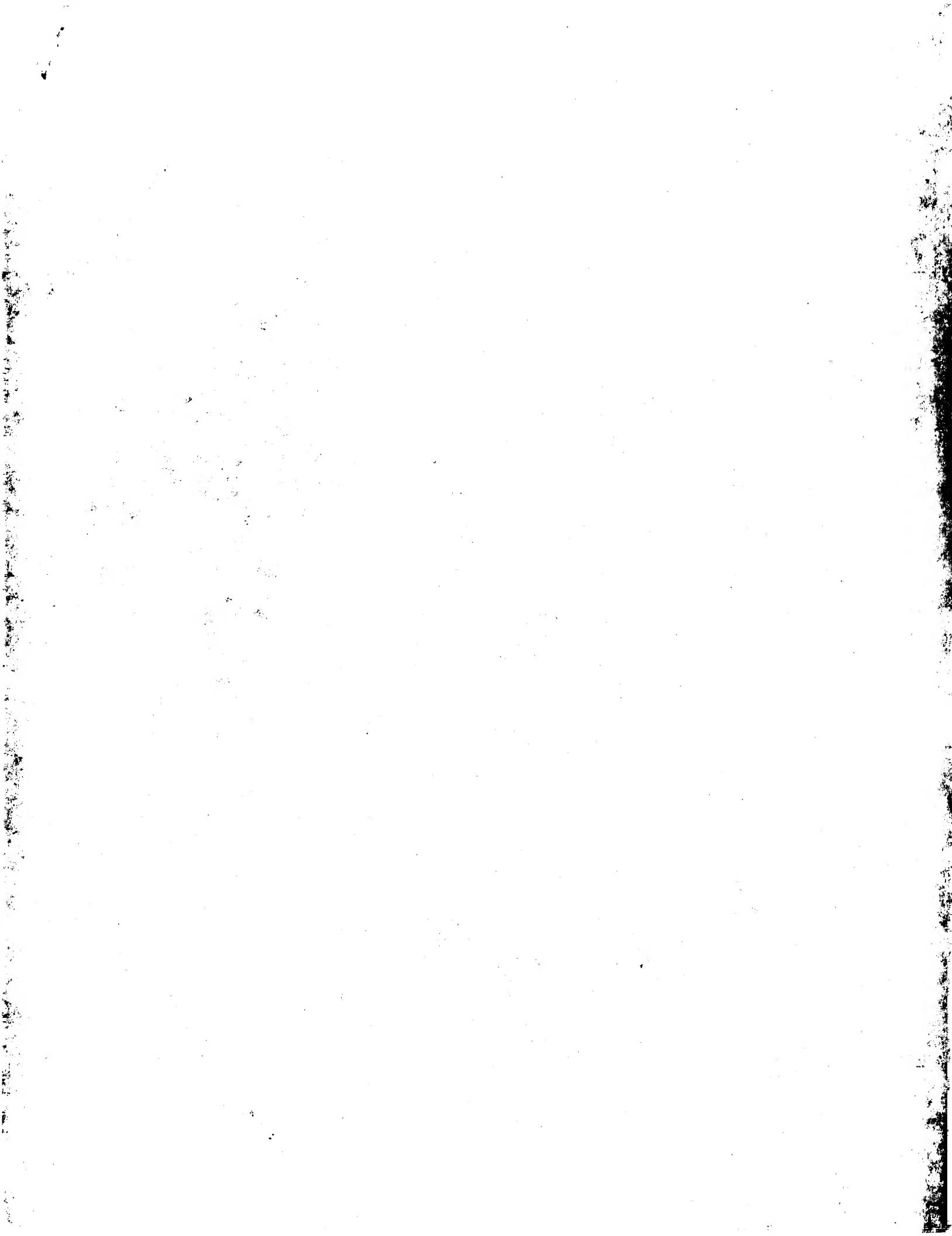
RESULT 9

Q9HZMO PRELIMINARY; PRT; 741 AA.

ID Q9HZMO
 AC Q9HZMO
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE PA2984;
 OS pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.

RC	STRAIN=ATCC 15692 / PAO1;	RESULT 11
RX	MEDLINE=2043733; PubMed=10984043;	Q9RXV3
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goitry L., Tolentino E., Westbroek-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Liim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J.J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V., "Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen."; Nature 406:959-964 (2000).	PRELIMINARY; PRT; 275 AA.
RA	EMBL: AE004724; AAC063721; -;	ID Q9RXV3
DR	InterPro; IPR001279; Blactmase-like.	AC Q9RXV3;
DR	InterPro; IPR004477; ComEC_N-term.	DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DR	Pfam: PF00753; lactamase_B; 1.	DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DR	TIGRFAMS; TIGR00360; ComEC_N-term; 1.	DE Hypothetical protein DR0203.
DR	TIGRFAMS; TIGR00361; ComEC_Rec2; 1.	GN DR0203.
KW	Hypothetical protein; Complete proteome.	OS Deinococcus radiodurans.
SQ	SEQUENCE 741 AA; 79979 MW; 2C17D3B56CEB98D0 CRC64;	OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales; Deinococcaceae; Deinococcus.
Query Match	Best Local Similarity 61.3%; Score 38; DB 16; Length 741; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	OX NCBI_TaxID=1299;
Oy	8 AWACGXA 14	RN [1]
Db	53 AWACGSA 59	RP SEQUENCE FROM N.A.
RESULT 10		RC STRAIN=R1;
ID	Q9A5NL	RX MEDLINE=2036896; PubMed=10567266;
AC	PRELIMINARY; PRT; 218 AA.	RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Ventker J.C., Fraser C.M.; "Genome sequence of the radioresistant bacterium <i>Deinococcus radiodurans</i> R1"; Science 286:1571-1577(1999).
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)	RL EMBL: AE001882; AAC097881; -.
DT	01-JUN-2002 (TREMBLrel. 21, last annotation update)	DR TIGR; DR0203; -.
DE	Transglycosylase, putative.	KW Hypothetical protein; Complete proteome.
GN	CC2416.	SQ SEQUENCE 275 AA; 29820 MW; F17E96929ED9EAA1 CRC64;
OS	Caulobacter crescentus.	Query Match Best Local Similarity 85.7%; Score 58; DB 16; Length 275; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OC	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;	RT radiodurans R1;
OC	Caulobacter.	RL Science 286:1571-1577(1999).
OX	NCBI_TaxID=155892;	DR EMBL: AE001882; AAC097881; -.
RN	[1]	DR TIGR; DR0203; -.
RP	SEQUENCE FROM N.A.	KW Hypothetical protein; Complete proteome.
RC	STRAIN=ATCC 15089 / CB15;	SQ SEQUENCE FROM N.A.
RX	MEDLINE=21173698; PubMed=11259647;	Query Match Best Local Similarity 59.7%; Score 37; DB 16; Length 275; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RA	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haff D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Yamada J., Embley M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M., "Complete genome sequence of <i>Caulobacter crescentus</i> "; Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).	RT putative hydrolase.
RA	EMBL: AE005911; AAC043871; -.	GN SC04552 OR SC016A.31.
DR	HSSP; P03810; IQSA.	OS Streptomyces coelicolor.
DR	TIGR; CC2416; -.	OC Bacteria; Firmicutes; Actinomycetidae; Actinomycetales; Streptomyceinae; Streptomyctaceae; Streptomyces.
DR	InteinPro; IPR000189; SLT_domain.	OX NCBI_TaxID=1902;
DR	Pfam; PF01464; SLN; 1.	RP SEQUENCE FROM N.A.
DR	PROSITE; PS00922; TRANSGLYCOSYLASE; 1.	RC STRAIN=A3(2) / MI45;
KW	Complete proteome.	RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Batman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Horrisby T., Howarth P., Huang C.-H., Kieser T., Lake L., Murphy L., Oliver K., O'Neill S., Rabbowitsch E., Rajandream M.A., Rutherford K., Rutledge S., Seeger K., Saunders D., Sharp S., Squares R., Taylor K., Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.; "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)"; Nature 417:141-147(2002).
SQ	SEQUENCE 218 AA; 23059 MW; 7E3106D6DFED6BB7 CRC64;	RL EMBL: AL078618; CAB44541; -.
Query Match	Best Local Similarity 59.7%; Score 37; DB 16; Length 218; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	KW Hydrolase.
Oy	8 AWACGXA 14	SQ SEQUENCE 334 AA; 33828 MW; 9789F3A066B587F1 CRC64;
Db	22 AWACGVA 28	



On protein - protein search, using sw model
Run on : March 25, 2003, 08:17:33 ; Search time 29.0309 Seconds
{without alignments})
45.805 Million cell updates/sec

Title: US-09-646-532B-1.
Perfect score: 48
Sequence: 1 SMPLNAAVKM 10
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq-101002:*

1: /SIDS1/gcdata/geneseq/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1990.DAT: *
11: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1991.DAT: *
12: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1992.DAT: *
13: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1993.DAT: *
14: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1994.DAT: *
15: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1995.DAT: *
16: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1996.DAT: *
17: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1997.DAT: *
18: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1998.DAT: *
19: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1999.DAT: *
20: /SIDS1/gcdata/geneseq/geneseq-emb1/AA2000.DAT: *
21: /SIDS1/gcdata/geneseq/geneseq-emb1/AA2001.DAT: *
22: /SIDS1/gcdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1
ID AAY42753
AAV42753 standard; peptide: 10 AA.
AC AAY42753:
XX DT 20-DEC-1999 (first entry)
XX DE wheat amyloplast ADP-glucose transporter peptide #1.
XX KW starch biosynthesis; amyloplast; ADP-glucose; transport; import;
KW PN optimisation; amylose; branching; chemical structure; transgenic plant;
KW OS triticum aestivum.
XX PN WO9947682-A1.
PD 23-SEP-1999.
XX PF 19-MAR-1999; 99WO-GB00728.
PR XX 20-MAR-1998; 98GB-0005939.
PA XX (UWMA-) UNIV VICTORIA MANCHESTER.
XX PT Eimes MJ, Tettlow IJ, Bowsher CG;
XX DR WPI; 1999-59097/50.
PT XX New transporter complex protein useful for modulating starch content in plants, especially useful in food production -
PT Streptococcus poly

Result No. Score Query Match Length DB ID Description
1 48 100.0 10 20 AAY42753 Wheat amyloplast A
2 48 100.0 10 20 AAY3931 ADP-gluco trans
3 35 72.9 87 23 ABP08699 Human ORFX protein
4 35 68.8 132 23 ABP07944 Human ORFX protein
5 33 68.8 36 22 AAM85180 Human immune/haema
6 32 66.7 56 23 ABP02748 Human ORFX protein
7 32 66.7 186 23 ABP27832 Streptococcus poly
8 32 66.7 323 23 Herbicidally active
9 32 66.7 402 23 ABP30120 Streptococcus poly
10 64.6 59 23 ABP26059

ALIGNMENTS

11 31 64.6 181 22 AAB68929 Neisseria meningit
12 31 64.6 185 22 AAU30414 Novel human secret
13 31 64.6 207 23 AAM50648 Arabidopsis BA:BA
14 31 64.6 315 23 ADF60886 Bacillus halodurans
15 31 64.6 363 18 AAW20581 Helicobacter pylor
16 31 64.6 363 18 AAW24695 H. pylori cell env
17 31 64.6 436 18 AAW55473 H. pylori ORF 06cp
18 31 64.6 436 18 AAW5543 H. pylori ORF 06cp
19 31 64.6 442 18 AAW55250 H. pylori ORF 05cp
20 31 64.6 509 20 AAY44020 Mouse Scavenger re
21 31 64.6 509 20 AAW93575 Murine scavenger r
22 31 64.6 509 23 AAB31525 Mouse ischaemic co
23 31 64.6 707 21 AAB32499 S. Enteritidis Mit
24 31 64.6 716 22 ABB66565 Drosophila melanog
25 31 64.6 823 19 AAW98507 H. pylori GHO 122
26 31 64.5 113 23 ABB50028 Extended human sec
27 30 62.5 168 22 AAG89204 Human secreted pro
28 30 62.5 195 22 AHO13539 Human secreted pro
29 30 62.5 72 23 ABB02945 Human AFP protein
30 30 62.5 76 22 AAU42162 Human AFP protein
31 30 62.5 113 23 ABB50028 Human AFP protein
32 30 62.5 168 22 AAG89204 Human AFP protein
33 30 62.5 195 22 AHO13539 Human AFP protein
34 30 62.5 236 22 AAGB1256 Human AFP protein
35 30 62.5 239 21 AAB01418 Mouse TANGO 232
36 30 62.5 258 22 ABB70394 Drosophila melanog
37 30 62.5 266 21 AAB14127 Bordetella pertussis
38 30 62.5 308 21 AAY88300 Human TANGO 187-3
39 30 62.5 315 23 ABB55045 Lactococcus lactis
40 30 62.5 319 21 AAY88296 Human TANGO 187-3
41 30 62.5 335 22 AAG91138 C. glutamicum prote
42 30 62.5 343 21 AY188281 Human TANGO 187-1/
43 30 62.5 343 22 AAGB1263 Human AFP protein
44 30 62.5 343 22 AAB88425 Human membrane or
45 30 62.5 354 21 AAY88297 Human TANGO 187-1/

PS	Claim 1; Page 3; 28pp; English.
XX	
CC	this sequence represents a wheat amyloplast ADP-glucose transporter peptide, #1. The wheat amyloplast membrane comprises at least two proteins; this sequence, along with peptides #2 (AY42754) and #3 (AY42755) are components of one of these proteins. Peptides #4-#7 (AY42756-AY42759) represent sequences within the second protein (AY42750). The sugar nucleotide ADP-glucose is the immediate substrate for starch synthesis, which occurs in the amyloplast; however, ADP-glucose is mainly synthesised outside the amyloplast in the cytoplasm. The ADP-glucose transporter is responsible for the import of ADP-glucose into the amyloplast and therefore plays a pivotal role in the regulation of starch synthesis. The transporter not only influences starch yield, but also quality as the starch synthases involved in amylose and amylopectin synthesis have different affinities for ADP-glucose. Variations in the chemical structure of starch are determined by the ratio of amylose to amylopectin, and by the degree of branching in amylopectin in the starch polymer. These variations can significantly alter the properties of starch. The ADP-glucose transporter complex is useful for generating transgenic plants in which the starch quality can be optimised for industrial applications in which starch is used. For example, transgenic plants which have an increased amylose content in starch are useful for production of starch with increased viscosity and gel strength, which prevents baked food going stale so quickly. Conversely, an increased amylopectin content in the starch produces waxy starch, useful as thickening agents in food and coatings.
CC	Sequence 10 AA;
CC	SQ
CC	Query Match 100.0%; Score 48; DB 20; Length 10; Best Local Similarity 100.0%; Pred. No. 0.00038; Mismatches 0; Matches 10; Conservative 0; Indels 0; Gaps 0;
Qy	1 SMPLNAAVMM 10
Db	1 SMPLNAAVMM 10
AC	AY39331;
AC	AY39331;
XX	01-DEC-1999 (first entry)
XX	DE ADP glucose transporter peptide 1.
XX	KW ADP glucose transporter; transform plant cell; wheat; starch production; gel strength; baked food.
XX	OS Triticum aestivum.
PN	W09947681-A1.
XX	PD 23-SEP-1999.
XX	PF 19-MAR-1999; 99WO-GB00727.
XX	PR 20-MAR-1998; 98GB-0005939.
XX	PA (UYMA-) UNIV VICTORIA MANCHESTER.
PT	Eimes MJ, Tetlow IJ, Bowsher CG;
XX	DR WPI; 1999-571841/48.
XX	PT ADP glucose transporter protein used for modifying plant starch production -
XX	PS Claim 1; Page 15; 26pp; English.

PR 14-JUL-2000; 200005-0218290.
 PR 26-JUL-2000; 200005-0220963.
 PR 14-AUG-2000; 200005-0224518.
 PR 14-AUG-2000; 200005-0224519.
 PR 14-AUG-2000; 200005-0225213.
 PR 14-AUG-2000; 200005-0225214.
 PR 14-AUG-2000; 200005-0225266.
 PR 14-AUG-2000; 200005-0225267.
 PR 14-AUG-2000; 200005-0225268.
 PR 14-AUG-2000; 200005-0225270.
 PR 14-AUG-2000; 200005-0225447.
 PR 14-AUG-2000; 200005-0225757.
 PR 14-AUG-2000; 200005-0227009.
 PR 18-AUG-2000; 200005-0226279.
 PR 22-AUG-2000; 200005-0226681.
 PR 22-AUG-2000; 200005-0226868.
 PR 22-AUG-2000; 200005-0227182.
 PR 23-AUG-2000; 200005-0227703.
 PR 30-AUG-2000; 200005-0228924.
 PR 01-SEP-2000; 200005-0229287.
 PR 01-SEP-2000; 200005-0229343.
 PR 01-SEP-2000; 200005-0229344.
 PR 01-SEP-2000; 200005-0229345.
 PR 05-SEP-2000; 200005-0229509.
 PR 05-SEP-2000; 200005-0229513.
 PR 06-SEP-2000; 200005-0230437.
 PR 08-SEP-2000; 200005-0230438.
 PR 08-SEP-2000; 200005-0231242.
 PR 08-SEP-2000; 200005-0231243.
 PR 08-SEP-2000; 200005-0231244.
 PR 08-SEP-2000; 200005-0231413.
 PR 08-SEP-2000; 200005-0231414.
 PR 08-SEP-2000; 200005-0232080.
 PR 12-SEP-2000; 200005-0231968.
 PR 14-SEP-2000; 200005-0232398.
 PR 14-SEP-2000; 200005-0232399.
 PR 14-SEP-2000; 200005-0232400.
 PR 14-SEP-2000; 200005-0232401.
 PR 14-SEP-2000; 200005-0233063.
 PR 14-SEP-2000; 200005-0233064.
 PR 14-SEP-2000; 200005-0233065.
 PR 21-SEP-2000; 200005-0234223.
 PR 21-SEP-2000; 200005-0234274.
 PR 25-SEP-2000; 200005-0234997.
 PR 25-SEP-2000; 200005-0234998.
 PR 27-SEP-2000; 200005-0235834.
 PR 27-SEP-2000; 200005-0235836.
 PR 29-SEP-2000; 200005-0236367.
 PR 29-SEP-2000; 200005-0236368.
 PR 29-SEP-2000; 200005-0236369.
 PR 29-SEP-2000; 200005-0236379.
 PR 02-OCT-2000; 200005-0236802.
 PR 02-OCT-2000; 200005-0237037.
 PR 02-OCT-2000; 200005-0237039.
 PR 02-OCT-2000; 200005-0237040.
 PR 13-OCT-2000; 200005-0239935.
 PR 13-OCT-2000; 200005-0239937.
 PR 20-OCT-2000; 200005-0240960.
 PR 20-OCT-2000; 200005-0241221.
 PR 20-OCT-2000; 200005-0241785.
 PR 20-OCT-2000; 200005-0241786.
 PR 20-OCT-2000; 200005-0241808.
 PR 20-OCT-2000; 200005-0241809.
 PR 20-OCT-2000; 200005-0241826.
 PR 01-NOV-2000; 200005-0244617.

PR 08-NOV-2000; 200005-0246474.
 PR 08-NOV-2000; 200005-0246475.
 PR 08-NOV-2000; 200005-0246476.
 PR 08-NOV-2000; 200005-0246477.
 PR 08-NOV-2000; 200005-0246478.
 PR 08-NOV-2000; 200005-0246523.
 PR 08-NOV-2000; 200005-0246524.
 PR 08-NOV-2000; 200005-0246525.
 PR 08-NOV-2000; 200005-0246526.
 PR 08-NOV-2000; 200005-0246527.
 PR 08-NOV-2000; 200005-0246528.
 PR 08-NOV-2000; 200005-0246532.
 PR 08-NOV-2000; 200005-0246609.
 PR 08-NOV-2000; 200005-0246610.
 PR 08-NOV-2000; 200005-0246611.
 PR 08-NOV-2000; 200005-0246613.
 PR 17-NOV-2000; 200005-024907.
 PR 17-NOV-2000; 200005-0249208.
 PR 17-NOV-2000; 200005-0249209.
 PR 17-NOV-2000; 200005-0249210.
 PR 17-NOV-2000; 200005-0249211.
 PR 17-NOV-2000; 200005-0249212.
 PR 17-NOV-2000; 200005-0249213.
 PR 17-NOV-2000; 200005-0249214.
 PR 17-NOV-2000; 200005-0249208.
 PR 17-NOV-2000; 200005-0249215.
 PR 17-NOV-2000; 200005-0249216.
 PR 17-NOV-2000; 200005-0249217.
 PR 17-NOV-2000; 200005-0249218.
 PR 17-NOV-2000; 200005-0249244.
 PR 17-NOV-2000; 200005-0249245.
 PR 17-NOV-2000; 200005-0249264.
 PR 17-NOV-2000; 200005-0249265.
 PR 17-NOV-2000; 200005-0249266.
 PR 17-NOV-2000; 200005-0249299.
 PR 17-NOV-2000; 200005-0249300.
 PR 01-DEC-2000; 200005-0249300.
 PR 01-DEC-2000; 200005-0249310.
 PR 05-DEC-2000; 200005-0251030.
 PR 05-DEC-2000; 200005-0251988.
 PR 05-DEC-2000; 200005-0256719.
 PR 06-DEC-2000; 200005-0251479.
 PR 08-DEC-2000; 200005-0251479.
 PR 08-DEC-2000; 200005-0251869.
 PR 08-DEC-2000; 200005-0251889.
 PR 08-DEC-2000; 200005-0251990.
 PR 11-DEC-2000; 200005-0254097.
 PR 05-JAN-2001; 200105-0259678.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM;
 DR WPI: 2001-483426/52.
 DR N-PSDB; AAK57961.
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
 XX Claim 11; SEQ ID NO 12773; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK8170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and Polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the

CC protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.

XX SQ Sequence 36 AA;

Query Match 68.9%; Score 33; DB 22; Length 36;
Best Local Similarity 66.3%; Pred. No. 4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 MPFLNAVKM 10
Db 2 MPFLNQTLKM 10

RESULT 6

ID ABP02748
ID ABP02748 standard; Protein; 56 AA.
AC ABP02748;
XX DT 24-JUN-2002 (first entry)
XX DE Human ORFX protein sequence SEQ ID NO:5478.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypertension; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune myasthenia gravis.
XX OS Homo sapiens.
XX PN WO200192523-A2.
XX PD 05-DRC-2001.
XX PR 29-MAY-2001; 2001WO-US10836.
XX PR 30-MAY-2000; 2000US-206132P.
XX PR 29-AUG-2000; 2000US-228716P.
(CURR-) CURAGEN CORP.
XX PI Shimkets RA, Leach MD;
XX DR WPI; 2002-106308/14.
DR N-PSDB; ABN18500.
XX PT Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders
PS Disclosure: SEQ ID 5478; 1037pp; English.

The present invention describes substantially purified human proteins (referred to as open reading frame, ORFX, where X is 1-1149 (see Table 1 in the specification). ABN15762 to ABN2252 encode the human ORFX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for treating or preventing a pathology associated with an ORFX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORFX-associated disorder. ORFX polynucleotide sequences can be used in gene therapy. ORFX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, varicose immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORFX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct Sequences.

XX SQ Sequence 56 AA;

Query Match 66.7%; Score 32; DB 23; length 56;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SMPLNAVKM 10
Db 24 NMPLNVAMP 33

RESULT 7

ID ABP27852
ID ABP27852 standard; Protein; 186 AA.
AC ABP27852;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 4880.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae; group A streptococcus; Streptococcus pyogenes; antibacterial; antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus agalactiae.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PR 29-OCT-2001; 2001WO-GB04789.
XX PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Mesignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX DR WPI; 2002-352536/38.
DR N-PSDB; ABN68483.

PT New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein -
PT XX PS Claim 1; Page 3650; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABW71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by

RESULT 10
 ABP26059
 ID ABP26059 standard; Protein: 59 AA.
 XX
 AC ABP26059;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 1294.
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PR 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-002633.
 PR 24-NOV-2000; 2000GB-002877.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masiognani V, Margarit Ros VI, Grandi G, Fraser C;
 PT Tettelin H;
 XX
 DR WPI; 2002-352535/38.
 DR N-PSDB; ABN66690.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS Claim 1; Page 326; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP3095) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus Pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6041-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 Sequence 59 AA;
 SQ Score 31; DB 23; Length 59;
 Best Local Similarity 71.4%; Pred. No. 20; Mismatches 5; Conservative
 Matches 5; Indels 0; Gaps 0;
 Query Match 64.6%; Score 31; DB 22; Length 181;
 Best Local Similarity 60.0%; Pred. No. 77; Mismatches 6; Conservative
 Matches 6; Indels 2; Gaps 0;
 QY 1 SMPLNAKWM 10
 :||| |||:
 Db 57 NMPLREAKVL 66
 RESULT 12
 AAU30414
 ID AAU30414 standard; Protein: 185 AA.
 XX
 AC AAU30414;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #905.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.

RESULT 11
 Oy 3 PLNAVK 9
 | :|||:
 Db 5 PINAAK 11

RESULT 11

XX
PR 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JRN-2001; 2001US-0770160.
XX
PA (HYSE-) HVSEQ INC.

XX
DR WPI; 2001-611725/70.
XX
PT Tang YT, Liu C, Drmanac RT;
PS
XX
CC Disclosure: FIG 12B; 78pp; English.

XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
Claim 20; Page 291; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
determining the presence of or predisposition to a disease associated
with altered levels of polypeptide. The polypeptides are also useful for
identifying agents (agonists and antagonists) that bind to them. Cells
expressing the proteins are useful for identifying a therapeutic agent
for use in treatment of a pathology related to aberrant expression or
physiological interactions of the polypeptides and cells genetically
engineered to express them are also useful for producing the proteins.
The proteins are useful in genetic vaccination, testing and
therapy, and can be used as nutritional supplements. They may be used to
increase stem cell proliferation; to regulate hematopoiesis; and in
bone, cartilage, tendon and/or nerve tissue growth or regeneration;
immune suppression and/or stimulation; as anti-inflammatory agents; and
in treatment of leukaemias. AU29510-Au33304 represent the amino acid
sequences of novel human secreted proteins of the invention.

Sequence 185 AA;

Query Match 64.6%; Score 31; DB 22; Length 185;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MPLNIAVKM 10
Dy 42 MPLNGYLMK 50

RESULT 13

AAM50648 ID ARM50648 standard; Protein; 207 AA.

AC AAM50648;
XX
DT 04-APR-2002 (first entry)

DE Arabidopsis BAL:BAP-like protein BAL.

XX
KW BAL; BAL:BAP-like protein; growth; homeostasis; thermotolerance;
KW transgenic plant; plant.
XX
OS Arabidopsis thaliana.
XX
PN WO200200697-A2.

XX
PD 03-JAN-2002.

XX
PF 25-JUN-2001; 2001WO-US20172.
PR 23-JUN-2000; 2000US-213863P.

XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Hua J, Grisafi P, Fink GR;
XX

DR WPI; 2002-139899/18.
XX
N-PSDB: ABA91260.

PT New phospholipid binding proteins and nucleic acids, useful for
modulating plant growth homeostasis, controlling cell expansion and
cell division, or producing plants where larger fruits and increased
biomass are desired -
XX
PT Disclosure: FIG 12B; 78pp; English.

CC The present sequence is that of the predicted protein product of
the BAL (BAL:Bap-like) protein gene of *Arabidopsis thaliana*. The
BAL protein is necessary for normal plant growth. The invention is
directed to isolated BON1, BON2, BON3, BAPI and BAL nucleic acids
(see ABA91256-56), which encode proteins (see AAM50641-48) that are
necessary for normal growth, controlling cell expansion and cell
division, and thereby affecting the size and rate at which the
plant grows when exposed to lower temperatures. Transgenic plants
are provided that are smaller than the wild-type as a result of
inhibition of BON1, BON3, BAPI and/or BAL, especially
angiosperms and gymnosperms, ornamental plants and turfgrass.
Transgenic plants are also provided that are larger than the
wild-type as a result of enhancement of BON1, BON2, BON3, BAPI
and/or BAL, especially crop plants and biomass plants. Modulation
of these genes provides increased yield, or growth at a higher
altitude or lower temperature.

Sequence 207 AA;

Query Match 64.6%; Score 31; DB 23; Length 207;
Best Local Similarity 62.3%; Pred. No. 91;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MPLNIAVK 9
Dy 64 MPINASVR 71

RESULT 14

ABP6086 ID ABP6086 standard; Protein; 315 AA.

AC ABP6086;
XX
DT 06-SEP-2002 (first entry)

DE *Bacillus halodurans* thioredoxin reductase SEQ ID NO:235.

XX
KW Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;
KW oil body; ophthalmological; antidiabetic; cytostatic; antipsoriatic;
KW vasoconstrictor; vulnerary; antibacterial; immunosuppressive; antiulcer;
KW food product; milk; wheat; oxidative stress; cataract; diabetes;
KW chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;
KW bronchiopulmonary disease; malignancy; reperfusion injury; sepsis;
KW gastro intestinal bleeding; intestinal bowel disease; ulcer;
KW gastro oesophageal reflux disease.
XX
OS *Bacillus halodurans*.
XX
PN WO200250289-A1.

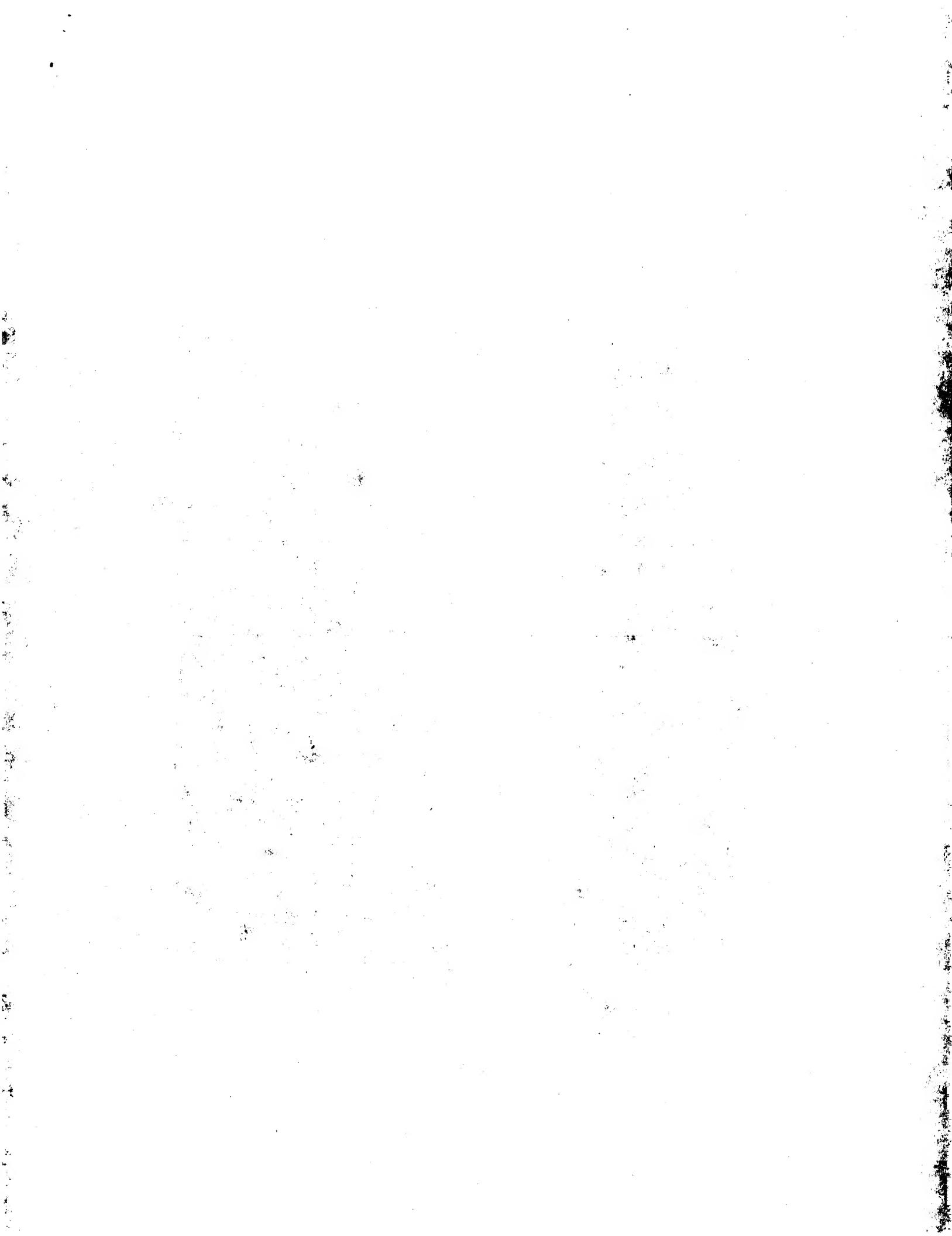
XX
PD 27-JUN-2002.

XX
PF 19-DEC-2001; 2001WO-US50340.

XX
PR 19-DEC-2000; 2000US-0743900.
PR 05-JUL-2001; 2001US-302865P.
PR 04-DEC-2001; 2001US-0006038.

XX
PA (SEM3-) SEMBIOYS GENETICS INC.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Van Roonen G, Deckers H, Heifetz PB, Briggs SR, Dalmia BK;

XX	PI	Del Val G, Zapachinski S, Moloney M;	XX
DR	WPI:	2002-508806/54.	XX
PT	Producing oil body associated with recombinant multimeric protein complex e.g. redox proteins and immunoglobulins comprises producing recombinant polypeptides capable of forming the complex in cells comprising oil bodies -	XX	PD 19-DEC-1996.
PT	XX	XX	XX
CC	The present invention describes a method (M1) for producing an oil body associated with a recombinant multimeric protein complex (MPC). M1 comprises producing in a cell comprising oil bodies a first and second recombinant polypeptide (P1, P2), where P1 is capable of associating with P2 to form the MPC and associating the complex with an occlusion body (OB) through an OB-targeting-protein capable of associating with OB and P1. M1 is useful for producing an oil body associated with a recombinant MPC. The oil bodies are further formulated for use in the preparation or a food product such as milk or wheat based food product, personal care product which reduces the oxidative stress on the surface area of the human body or used to lighten the skin, or a pharmaceutical composition used to treat chronic obstructive pulmonary disease (COPD), cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis, malignancies, reperfusion injury, wound healing, sepsis, gastro-intestinal (GIN) bleeding, intestinal bowel disease (IBD), ulcers, GERD (gastric oesophageal reflux disease). ABN89569 to ABN9593 and ABP60677 to ABP6064 represent sequence given in the exemplification of the present invention.	XX	PR 06-JUN-1996; 96WO-US09122.
CC	XX	XX	PR 01-APR-1996; 96US-0630-05.
CC	XX	XX	PR 07-JUN-1995; 95US-0487032.
PA	(ASTR) ASTRA AB.	XX	XX
PA	Berglindh OT, Smith D, Mellgaard BL;	XX	XX
PA	DR N-PSDB; AAN67834.	XX	XX
PT	Helicobacter pylori nucleic acid sequences and related polypeptide(s) - useful for vaccines to treat or prevent H. pylori infection, and to detect Helicobacter	XX	PT 1997-052306/05.
PT	H. pylori infection or to identify H. pylori poly peptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors.	XX	PT 1997-052306/05.
CC	The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or reported membrane proteins. Identified and determined the sequences of interest. Particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.	XX	CC 06-JUN-1996; 96WO-US09122.
CC	A 100 microg dose of the protein was given to female SPF BALB/C mice infected with a H. pylori type 2 strain. The mice were also given 10 microg of cholera toxin as adjuvant. Omeprazole (400 micromol/kg) was given orally 3-5 hours prior to immunisation to protect the protein from degradation. The mice were sacrificed 2-4 weeks after final immunisation and their serum antibody titre determined to be about 400, compared to about 350 for a Lacto control.	XX	CC 01-APR-1996; 96US-0630-05.
SQ	Sequence 315 AA:	XX	CC 07-JUN-1995; 95US-0487032.
Query Match	Score 31; DB 23; Length 315;	XX	XX
Best Local Similarity	75.0%;	XX	XX
Matches	Pred. No. 1.5e+02;	XX	XX
Qy	2 MPLNAVK 9	XX	XX
Db	243 LPNEAVK 250	XX	XX
DT	04-JUN-1997 (first entry)	XX	XX
XX	DE Helicobacter pylori cell envelope protein, 978477.aa.	XX	XX
XX	Cell envelope; vaccine; prevention; treatment; infection; identification; binding compound; bacterium; life cycle; activator; bacteria; inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.	XX	XX
XX	KW	XX	XX
OS	Helicobacter pylori.	XX	XX
FH	Key Location/Qualifiers	XX	XX
FT	Misc-difference 11 /note= "unknown"	XX	XX
FT	Misc-difference 337 /note= "unknown"	XX	XX
FT	Misc-difference 339 /note= "unknown"	XX	XX
FT	Misc-difference 356 /note= "unknown"	XX	XX
FT	Misc-difference 360 /note= "unknown"	XX	XX
FT	Misc-difference 363 /note= "unknown"	XX	XX
PN	WO96640893-A1.	XX	XX
Search completed: March 25, 2003, 08:19:40	;	;	;
Job time : 31.0909 secs	;	;	;



GenCore version 5.1.4-p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 08:21:42 ; Search time 10.6051 Seconds

(without alignments)
 27.742 Million cell updates/sec

Title: US-09-646-532B-1
 Perfect score: 48
 Sequence: 1 SMPILNAAVKM 10
 Scoring table: BLOSUM62
 Gapext 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa5A_COMB_pep: *
 2: /cgn2_6/ptodata/2/iaa5B_COMB_pep: *
 3: /cgn2_6/ptodata/2/iaa6A_COMB_pep: *
 4: /cgn2_6/ptodata/2/iaa6B_COMB_pep: *
 5: /cgn2_6/ptodata/2/iaaPCTUS_COMB_pep: *
 6: /cgn2_6/ptodata/2/iae/backfilled1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	64.6	509	2	US-09-559-505-4
2	31	64.6	509	2	US-08-749-907-4
3	31	64.6	509	4	US-09-241-581B-8
4	31	64.6	509	5	PCM-US955721-8
5	30	62.5	291	2	US-08-701-19A-22
6	30	62.5	410	4	US-09-134-001C-4135
7	30	62.5	1135	2	US-08-469-537A-97
8	30	62.5	1138	1	US-08-323-474-8
9	30	62.5	1138	2	US-08-459-537A-98
10	30	62.5	1138	2	US-08-220-405-5
11	29	60.4	210	4	US-09-247-155-121
12	29	60.4	263	2	US-08-832-690-3
13	29	60.4	264	3	US-08-924-570-2
14	29	60.4	528	3	US-08-680-988-6
15	28	58.3	169	4	US-09-311-620B-20
16	28	58.3	236	1	US-08-266-570B-4
17	28	58.3	236	1	US-08-266-570B-6
18	28	58.3	236	1	US-08-266-570B-14
19	28	58.3	236	2	US-08-757-046A-14
20	28	58.3	236	3	US-09-447-208-14
21	28	58.3	236	4	US-09-135-988-14
22	28	58.3	236	4	US-09-277-716-14
23	28	58.3	236	4	US-08-597-274A-14
24	28	58.3	236	4	US-09-909-909-14
25	28	58.3	236	4	US-09-161B-14
26	28	58.3	236	4	US-08-990-103-14
27	28	58.3	236	4	US-09-154-802-1
28	58.3				

RESULT 1
 US-08-559-505-4
 Sequence 4, Application US/08559505
 Patent No. 592533
 GENERAL INFORMATION:
 APPLICANT: Monty Krleger, Susan L. Action, Attilio Rigotti, Helen H.
 TITLE OF INVENTION: METHODS FOR MODULATION OF LIPID UPTAKE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patrea L. Pabst
 STREET: 2800 One Atlantic Center
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30309-3450
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/559, 505
 FILING DATE: 29-NOV-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Patrea L.
 REGISTRATION NUMBER: 31-284
 REFERENCE/DOCKET NUMBER: MITT150
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..509
 OTHER INFORMATION: /Function = "Amino acid sequence for the murine Scavenger Receptor Class B1."

OTHER INFORMATION: murine Scavenger Receptor Class B1.
 US-08-559-505-4
 Query Match 64.6%; Score 31; DB 2; Length 509;
 Best Local Similarity 55.6%; Pred. No. 1e02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MPLNAVKM 10
 :|:|:||
 Db 380 IPMNCVKM 388

RESULT 2
 US-08-749-907-4 Application US/08749907
 Sequence 4, Application US/08749907
 Patent No. 5962322
 GENERAL INFORMATION:
 APPLICANT: Monty Krieger, Artilio Rigotti, and Karen Kozarsky
 TITLE OF INVENTION: METHODS FOR MODULATION OF CHOLESTEROL
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patrea L. Pabst, Arnall Golden & Gregory
 STREET: 2800 One Atlantic Center
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30309-3450
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 MEDIUM TYPE: Floppy disk
 SOFTWARE: PatientIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/241,581B
 FILING DATE: 02-Feb-1999
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: MIT6620
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: MIT7538
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: misc. feature
 LOCATION: 1..509
 OTHER INFORMATION: /Function = "Amino acid sequence for
 the murine Scavenger Receptor Class BI."
 SEQUENCE DESCRIPTION: SEQ ID NO: 8
 US-09-241-581B-8

Query Match 64.6%; Score 31; DB 4; Length 509;
 Best Local Similarity 55.6%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;
 Matches 5; Conservative 3; Mismatches 1;
 Qy 2 MPLNAVKM 10
 Db 380 IPMNCVKM 388

RESULT 4
 PCT-US95-07721-8
 Sequence 8, Application PC/TUS9507721
 GENERAL INFORMATION:
 APPLICANT: Massachusetts Institute of Technology
 TITLE OF INVENTION: Class BI and CI Scavenger Receptors
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patrea L. Pabst
 STREET: 2800 One Atlantic Center
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30309-3450
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 MEDIUM TYPE: Floppy disk
 SOFTWARE: PatientIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/07721
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: MIT6620
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 8:
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 APPLICANT: Massachusetts Institute of Technology
 TITLE OF INVENTION: Class BI and CI Scavenger Receptors
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patrea L. Pabst
 STREET: 2800 One Atlantic Center
 CITY: Atlanta
 STATE: Georgia

SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: misc. feature
 LOCATION: 1..509
 OTHER INFORMATION: /Function = "Amino acid sequence for the murine Scavenger Receptor Class B1."
 PCT-US95-07721-8

Query Match 64.6%; Score 31; DB 5; Length 509;
 Best Local Similarity 55.6%; Pred. No. 1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MPLNAAVM 10
 Db 380 IPMNCVVKM 388

RESULT 5
 US-08-701-191A-22
 Sequence 22, Application US/08701191A
 Patent No. 5942428

GENERAL INFORMATION:
 APPLICANT: Moosa Mohammadji, Joseph Schlessinger,
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN RECEPTOR TYROSINE KINASE NUMBER OF SEQUENCES: 41
 CURRENT APPLICATION DATA:
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C., DOS 5.0
 SOFTWARE: FastSEQ for Windows 2.0

PRIOR APPLICATION NUMBER: US/08-701,191A
 FILING DATE: August 21, 1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 227/088
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 291 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-701-191A-22

Query Match 62.5%; Score 30; DB 4; Length 410;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SMPLNAAV 8
 Db 11 TMPLNAVM 18

RESULT 6
 US-09-134-001C-4135
 Sequence 4135, Application US/09134001C
 Patent No. 6380370

GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134, 001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064, 964
 PRIOR FILING DATE: 1997-11-08
 PRIORITY NUMBER: US 60/055, 779
 PRIORITY FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 4135
 LENGTH: 410
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis

RESULT 7
 US-08-469-537A-97
 Sequence 97, Application US/08469537A
 Patent No. 5843749

GENERAL INFORMATION:
 APPLICANT: Maizonpierre, et al.
 TITLE OF INVENTION: ERK AND ROR TYROSINE KINASES
 TITLE OF INVENTION: ERK AND ROR TYROSINE
 NUMBER OF SEQUENCES: 107
 CURRENT APPLICATION DATA:
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 STREET: 777 Old Saw Mill River Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: U.S.A.
 ZIP: 10591

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 2.0

PRIOR APPLICATION NUMBER:
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 APPLICATION NUMBER: USSN 08/406, 247
 FILING DATE: 17-MAR-1995
 APPLICATION NUMBER: USSN 08/144, 992
 FILING DATE: 28-OCT-1993
 APPLICATION NUMBER: USSN 07/736, 559
 FILING DATE: 26-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kempler, Ph.D., Gail M
 REGISTRATION NUMBER: 32,143
 REFERENCE/DOCKET NUMBER: REG 070C
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-5400
 TELEFAX: 914-345-7721
 TELEX:
 INFORMATION FOR SEQ ID NO: 97:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1135 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-469-537A-97

Query Match 62.5%; Score 30; DB 2; Length 1135;
 Best Local Similarity 55.6%; Pred. No. 4.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MPLNAVKM 10
 : :|||:
 Db 860 LKMNNAIKM 868

RESULT 8
 Sequence 8, Application US/08323474
 Patent No. 5447860
 GENERAL INFORMATION:
 APPLICANT: Ziegler, Steven F.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98101
 COMPUTER READABLE FORM:
 COMPUTER TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatientIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/323,474
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USN 08/406,247
 FILING DATE: 17-MAR-1995
 APPLICATION NUMBER: USN 08/714,992
 FILING DATE: 28-OCT-1993
 APPLICATION NUMBER: USN 07/736,559
 FILING DATE: 26-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kempler, Ph.D., Gail M
 REGISTRATION NUMBER: 32,143
 REFERENCE/DOCKET NUMBER: REG 070C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 914-345-7400
 TELEFAX: 914-345-7721
 INFORMATION FOR SEQ ID NO: 98:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1138 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-469-537A-98

Query Match 62.5%; Score 30; DB 2; Length 1138;
 Best Local Similarity 55.6%; Pred. No. 4.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MPLNAVKM 10
 : :|||:
 Db 863 LKMNNAIKM 871

RESULT 10
 Sequence 5, Application US/08220240A
 Patent No. 595591
 GENERAL INFORMATION:
 APPLICANT: Alitalo, Kari
 APPLICANT: Matikainen, Maria-Terttu
 APPLICANT: Partanen, Juhu
 APPLICANT: Makela, Tomi
 APPLICANT: Korhonen, Jaana
 TITLE OF INVENTION: ANTIODIES RECOGNIZING TIE RECEPTOR
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:

Query Match 62.5%; Score 30; DB 1; Length 1138;
 Best Local Similarity 55.6%; Pred. No. 4.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MPLNAVKM 10
 : :|||:
 Db 863 LKMNNAIKM 871

Sequence 2, Application US/08924570A
 Patent No. 6093600
 GENERAL INFORMATION:
 APPLICANT: Reiter, Robert
 APPLICANT: Witte, Owen
 APPLICANT: Sawyer, Charles
 TITLE OF INVENTION: E25a PROTEIN, METHODS FOR
 TITLE OF INVENTION: PRODUCTION AND USE THEREOF
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Mandel & Adriano
 STREET: 725 Main Street
 CITY: Half Moon Bay
 STATE: CA
 COUNTRY: USA
 ZIP: 94019
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/924,570A
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT US97/15606
 FILING DATE: 05-SEP-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Adriano, Sarah B.
 REGISTRATION NUMBER: 34,470
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 213-254-1940
 TELEFAX: 213-258-5580
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 264 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-924-570A-2

RESULT 14
 US-08-988-988-6
 Sequence 6, Application US/08688988B
 Patent No. 609545
 GENERAL INFORMATION:
 APPLICANT: Lefebvre, Daniel D.
 APPLICANT: Malboobi, Mohammad A.
 TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
 FILE REFERENCE: PPL96-03
 CURRENT APPLICATION NUMBER: US/08/688,988B
 CURRENT FILING DATE: 1996-07-31
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 6
 LENGTH: 528
 TYPE: PRT
 ORGANISM: Arabidopsis Thaliana
 US-08-688-988-6

Query Match 60.4%; Score 29; DB 3; Length 528;
 Best Local Similarity 75.0%; Pred. No. 2.8e+02; Matches 6; Conservative 1; MisMatches 1; Indels 0; Gaps 0;
 QY 1 SMPNAAV 8
 Db 388 SMPLTAAL 395

RESULT 15
 US-09-311-626B-20
 Sequence 20, Application US/09311626B
 Patent No. 6399347
 GENERAL INFORMATION:
 APPLICANT: Jorgensen, Per Lina
 APPLICANT: Schnorr, Kirk
 APPLICANT: Andersen, Leine NO. 6399347bce
 APPLICANT: Schulein, Martin
 APPLICANT: Outrump, Helle
 FILE REFERENCE: 5572-204-US
 CURRENT APPLICATION NUMBER: US/09/311,626B
 CURRENT FILING DATE: 1999-05-13
 PRIORITY NUMBER: 060898
 PRIORITY FILING DATE: 1998-05-01
 PRIORITY APPLICATION NUMBER: 60/084,358
 PRIORITY FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 61
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 20
 LENGTH: 169
 TYPE: PRT
 ORGANISM: Bacillus agaradhaerens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: 147
 OTHER INFORMATION: Xaa = any amino acid
 US-09-311-626B-20

Query Match 58.3%; Score 28; DB 4; Length 169;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02; Matches 5; Conservative 3; MisMatches 0; Indels 0; Gaps 0;
 QY 2 MPLNAAVK 9
 Db 17 MPLHAAMR 24

Search completed: March 25, 2003, 08:22:58
 Job time : 11.6061 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.4-p5-4578

DM protein - protein search, using sw model

run on: March 25, 2003, 08:22:27 ; Search time 9,393.94 Seconds
 (without alignments)
 56,911 Million cell updates/sec

Title: us-09-646-532B-1
Perfect score: 48
Sequence: 1 SMPLNAAVKKM 10

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

221133 seqs, 53462247 residues

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

database :

Published Applications AA: *

- 1: /cgn2_6/ptodata/2/pubpaa/us08_NEWPUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/pep_NEWPUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/us06_NEWPUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/us05_NEWPUB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/us07_NEWPUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/us07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/pctus_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/us08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/us09_NEWPUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/us09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/us10_NEWPUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/us10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/us60_NEWPUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/us60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result NO.	Score	Query Match Length	DB ID	Description
1	31	64.6	207	9 US-09-891-139A-11 Sequence 11, Appl
2	31	64.6	509	10 US-09-891-139A-11 Sequence 4, Appl
3	31	64.6	823	9 US-09-895-913A-244 Sequence 244, App
4	30	62.5	168	10 US-09-871-324 Sequence 324, App
5	30	62.5	239	10 US-09-895-858-18 Sequence 18, Appl
6	30	62.5	335	9 US-09-738-626-4892 Sequence 4892, Ap
7	29	60.4	18	9 US-09-820-0969-5 Sequence 5, Appl
8	29	60.4	18	10 US-09-820-296-5 Sequence 5, Appl
9	29	60.4	263	9 US-10-174-590-184 Sequence 184, App
10	29	60.4	263	9 US-10-176-757-184 Sequence 184, App
11	29	60.4	263	9 US-10-176-913-184 Sequence 184, App
12	29	60.4	263	9 US-10-176-552-184 Sequence 184, App
13	29	60.4	263	9 US-10-175-738-184 Sequence 184, App
14	29	60.4	263	9 US-10-175-752-184 Sequence 184, App
15	29	60.4	263	9 US-10-176-482-184 Sequence 184, App
16	29	60.4	263	9 US-10-176-757-184 Sequence 184, App
17	29	60.4	263	9 US-10-176-913-184 Sequence 184, App
18	29	60.4	263	9 US-10-180-557-184 Sequence 184, App
19	29	60.4	263	9 US-10-180-557-184 Sequence 184, App

ALIGNMENTS

RESULT 1
US-09-891-139A-11
 Sequence 11, Application US/09891139A ; Publication No. US20020194639A1 ; GENERAL INFORMATION: APPLICANT: Hu, Jian ; APPLICANT: Gribaffi, Paula ; APPLICANT: Fink, Gerald R. ; TITLE OF INVENTION: Bonsai, A Phospholipid Bilayer Model for the Study of Membrane Proteins ; TITLE OF INVENTION: IS Required For Thermal, File Reference: 0399-009-001 CURRENT APPLICATION NUMBER: US/09/891,139A ; CURRENT FILING DATE: 2002-06-24 ; PRIORITY APPLICATION NUMBER: US 60/213,863 ; NUMBER OF SEQ ID NOS: 12 ; SOFTWARE: FastSEQ for Windows Version 4.0 ; SEQ ID NO 11 ; LENGTH: 207 ; TYPE: PRT ; ORGANISM: Arabidopsis Thaliana ; FEATURE: PEPTIDE ; NAME/KEY: PEPTIDE ; LOCATION: (1)..(207) ; OTHER INFORMATION: BAIL Protein ; US-09-891-139A-11

Query Match 64.6%; Score 31; DB Best Local Similarity 62.5%; Pred. No. 51; Matches 5; Conservative 3; Mismatches 0

RESULT 2
US-09-148-012-4
 QY 2 MPLNAAV 9
 ||:||:||:
 Db 64 MPINASVR 71

; Sequence 4, Application US/09148012 ; Patent No. US20020090040A1 ; GENERAL INFORMATION: APPLICANT: Krieger, Monty

SUMMARIES

ALIGNMENT

TITLE OF INVENTION: SR-BI Antagonist And Use Thereof As Contraceptives And FILE REFERENCE: MIT-150CIP2

CURRENT APPLICATION NUMBER: US/09/7148, 012

EARLIER APPLICATION NUMBER: 60/057, 943

EARLIER FILING DATE: 1997-09-05

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 509

TYPE: PRT

ORGANISM: Mouse

US-09-148-012-4

RESULT 3

Query Match Score 31; DB 10; Length 509;

Best Local Similarity 55.6%; Pred. No. 1.4e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 MPLNRAVKM 10

Db 380 IPMNCVYKM 388

US-09-895-913A-244

Sequence 244, Application US/09895913A

PATENT NO.: US20020160456A1

GENERAL INFORMATION:

APPLICANT: Kleanthou, Harold

APPLICANT: Al-Garawi, Amal

APPLICANT: Miller, Charles

APPLICANT: Tomb, Jean Francois

APPLICANT: Oomen, Raymond P.

TITLE OF INVENTION: Identification of Polynucleotides

FILE REFERENCE: 06132/043002

CURRENT APPLICATION NUMBER: US/09/895, 913A

CURRENT FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 08/881, 227

PRIOR FILING DATE: 1997-06-24

NUMBER OF SEQ ID NOS: 368

SOFTWARE: FASSEQ FOR Windows Version 4.0

SEQ ID NO 244

LENGTH: 823

TYPE: PRT

ORGANISM: Helicobacter pylori

US-09-895-913A-244

RESULT 4

Query Match Score 31; DB 9; Length 823;

Best Local Similarity 55.6%; Pred. No. 2.3e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SMPNLAALK 9

Db 639 SLPNLNSIK 647

US-09-731-872-324

Sequence 324, Application US/09731872

PATENT NO.: US20020102604A1

GENERAL INFORMATION:

APPLICANT: Dumas, Mine Edwards, Jean Baptiste

APPLICANT: Bougueret, Lydie

APPLICANT: Jobert, Severin

TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS

FILE REFERENCE: 78 US3, REG

CURRENT APPLICATION NUMBER: US/09/731, 872

PRIOR APPLICATION NUMBER: US 60/169, 629

PRIOR FILING DATE: 1999-12-08

US-09-796-858-18

Sequence 18, Application US/09796858

PATENT NO.: US20020055139A1

GENERAL INFORMATION:

APPLICANT: Holtzmann, Douglas

TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, AND OTHER USES

FILE REFERENCE: 7853-226-999

CURRENT APPLICATION NUMBER: US/09/796, 858

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 09/223, 094

PRIOR FILING DATE: 1998-12-30

PRIOR APPLICATION NUMBER: 09/223, 546

PRIOR FILING DATE: 1998-12-30

PRIOR APPLICATION NUMBER: 09/224, 246

PRIOR FILING DATE: 1998-12-30

PRIOR APPLICATION NUMBER: 09/312, 359

PRIOR FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 09/336, 536

PRIOR FILING DATE: 1999-06-18

PRIOR APPLICATION NUMBER: 09/342, 687

PRIOR FILING DATE: 1999-06-29

PRIOR APPLICATION NUMBER: 09/399, 723

PRIOR FILING DATE: 1999-09-20

PRIOR APPLICATION NUMBER: 09/471, 179

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: 09/474, 071

PRIOR FILING DATE: 1999-12-29

PRIOR APPLICATION NUMBER: 09/474, 072

PRIOR FILING DATE: 1999-12-29

PRIOR APPLICATION NUMBER: 09/572, 002

PRIOR FILING DATE: 2000-05-14

PRIOR APPLICATION NUMBER: 09/597, 993

PRIOR FILING DATE: 2000-06-12

PRIOR APPLICATION NUMBER: 09/599, 596

PRIOR FILING DATE: 2000-06-22

PRIOR APPLICATION NUMBER: 09/606, 565

PRIOR FILING DATE: 2000-06-29

PRIOR APPLICATION NUMBER: 09/365, 164

PRIOR FILING DATE: 1999-07-30

PRIOR APPLICATION NUMBER: 09/630, 334

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: 09/665, 666

PRIOR FILING DATE: 2000-09-20

NUMBER OF SEQ ID NOS: 50

SEQ ID NO 18

LENGTH: 239

TYPE: PRT

ORGANISM: Homo sapiens

US-09-796-858-18

Query Match 62.5%; Score 30; DB 10; Length 239;
 Best Local Similarity 66.7%; Pred. No. 96; Mismatches 2;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SMPDNAVK 9
 Db 128 SAPLNATVR 136

RESULT 6 US-09-738-626-4892
 Sequence 4892, Application USA/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NNOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 4892
 LENGTH: 335
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4892

Query Match 62.5%; Score 30; DB 9; Length 335;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MPLNAA 7
 Db 164 MPLNAA 169

RESULT 7 US-09-820-036B-5
 Sequence 5, Application US/09820096B
 Publication No. US20030022168A1
 GENERAL INFORMATION:
 APPLICANT: Davahara, Hiroyuki
 APPLICANT: Lewis, Laurence
 APPLICANT: Lewis, No. US20030022168A1
 TITLE OF INVENTION: ARYL PROPEENAL DOUBLE BOND REDUCTASE
 FILE REFERENCE: WSUR-1-17233
 CURRENT APPLICATION NUMBER: US/09/820, 096B
 CURRENT FILING DATE: 2001-03-27
 PRIOR APPLICATION NUMBER: US 60/192, 265
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 5
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Pinus Taeda

RESULT 8 US-09-820-296-5
 Sequence 5, Application US/09820296
 Publication No. US20030048424A1
 GENERAL INFORMATION:
 APPLICANT: Kasahara, Hiroyuki
 APPLICANT: Davin, Laurence
 APPLICANT: Lewis, No. US20010048424A1
 TITLE OF INVENTION: ARYL PROPENAL DOUBLE BOND REDUCTASE
 FILE REFERENCE: WSUR-1-17233
 CURRENT APPLICATION NUMBER: US/09/820, 296
 CURRENT FILING DATE: 2001-05-29
 PRIOR APPLICATION NUMBER: US 60/192, 266
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 5
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Pinus Taeda
 US-09-820-296-5

Query Match 60.4%; Score 29; DB 9; Length 18;
 Best Local Similarity 55.6%; Pred. No. 8.9; Mismatches 2;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SMPDNAVK 9
 Db 10 SFPLNQAIR 18

RESULT 9 US-10-174-590-184
 Sequence 184, Application US/10174590
 Publication No. US20030008352A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Simith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C42
 CURRENT APPLICATION NUMBER: US/10/174,590
 CURRENT FILING DATE: 2003-06-18
 Prior application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 184
 LENGTH: 263
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-174-590-184

Query Match 60.4%; Score 29; DB 9; Length 263;
 Best Local Similarity 55.6%; Pred. No. 1.7e+02; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SMPDNAVK 9
 Db 10 SFPLNQAIR 18

Qy 2 MPLNAAVKM 10
US-10-176-758-184
; Sequence 184, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; PRIOR Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-758-184

Query Match 60.4%; Score 29; DB 9; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.7e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPLNAAVKM 10
Db 163 MPLNTSIVM 171

RESULT 11
US-10-175-737-184
; Sequence 184, Application US/10175737
; Publication No. US2003013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC7
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-17
; PRIOR Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-175-737-184

Query Match 60.4%; Score 29; DB 9; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.7e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPLNAAVKM 10
Db 163 MPLNTSIVM 171

RESULT 13
US-10-175-738-184
; Sequence 184, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; PRIOR Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-175-737-184

Query Match 60.4%; Score 29; DB 9; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.7e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPLNAAVKM 10
Db 163 MPLNTSIVM 171

```

; ORGANISM: Homo Sapien ; SEQ ID NO 184
; US-10-175-738-184 ; LENGTH: 263
; Query Match 60.4%; Score 29; DB 9; Length 263;
; Best Local Similarity 55.6%; Pred. No. 1.7e+02; ; TYPE: PRT
; Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
; Qy 2 MPLNAAVKM 10 ; Db 163 MPLNTSIVM 171
; Db 163 MPLNTSIVM 171

RESULT 14
US-10-175-752-184
; Sequence 184, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC60
; CURRENT APPLICATION NUMBER: US/10/175, 752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-175-752-184

Query Match 60.4%; Score 29; DB 9; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.7e+02; ; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 15
US-10-176-482-184
; Sequence 184, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC70
; CURRENT APPLICATION NUMBER: US/10/176, 482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-482-184

Query Match 60.4%; Score 29; DB 9; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.7e+02; ; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

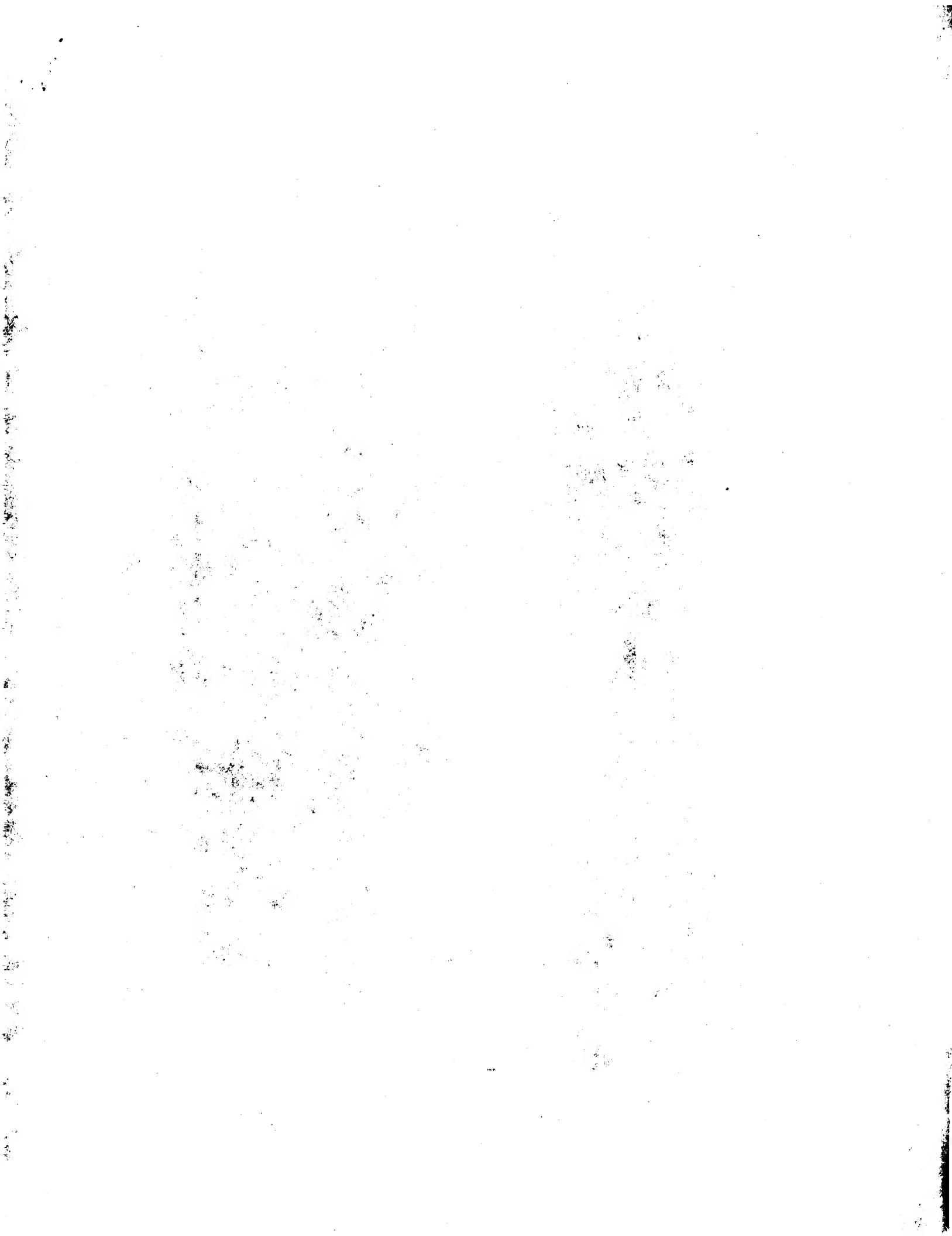
```

Search completed: March 25, 2003, 08:23:38
 Job time : 11.3939 secs

```

; SEQ ID NO 184
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-482-184
; Query Match 60.4%; Score 29; DB 9; Length 263;
; Best Local Similarity 55.6%; Pred. No. 1.7e+02; ; Indels 0; Gaps 0;
; Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
; Qy 2 MPLNAAVKM 10
; Db 163 MPLNTSIVM 171
; Db 163 MPLNTSIVM 171

```



A; Reference number: A89758;	N; Accession: K99392	2	MPLNNAVK	9
A; Status: preliminary			:	
A; Molecule type: DNA		74	MPLNAVVR	81
A; Residues: 1-459 <KUR>				
A; Cross-references: GB:BA000018; PID:913702126; PIDN:BAB43418.1; GSPDB:GN00149				
C; Genetics:				
A; Gene: SA2117				
RESULT 3				
Query Match	69.8%	Score 33;	DB 2;	Length 459;
Best Local Similarity	75.0%	Pred. No. 32;		
Matches 6;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	8		
Db 56	SMPLNVAI	63		
RESULT 4				
Query Match	69.8%	Score 33;	DB 2;	Length 459;
Best Local Similarity	75.0%	Pred. No. 32;		
Matches 6;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	10		
Db 138	SLPLNADPVE	147		
RESULT 5				
Query Match	69.8%	Score 33;	DB 2;	Length 459;
Best Local Similarity	75.0%	Pred. No. 32;		
Matches 6;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	8		
Db 56	SMPLNVAI	63		
RESULT 3				
translation initiation inhibitor [imported] - <i>Brucella melitensis</i> (strain 16M)				
C; Species: <i>Brucella melitensis</i>				
C; Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002				
C; Accession: AF3354				
R; DelVecchio, V.; Kapatral, V.; Redkar, R.J.; Patra, G.; Muier, C.; Los, T.; Ivanova, N.; Mazur, M.; Goitsman, E.; Selkov, E.; Elzer, P.H.; Hagiwara, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 433-448, 2002				
A; Title: The genome sequence of the facultative intracellular pathogen <i>Brucella melitensis</i> R; Skala, J.; Nawrocki, R.; Goffeau, A.				
A; Reference number: AD3252; PMID:11756688				
A; Accession: AF3354				
A; Status: preliminary				
A; Molecule type: DNA				
A; Residues: 1-154 <KUR>				
A; Cross-references: GB:AE008917; PIDN:AAL52001.1; PID:917982764; GSPDB:GN00190				
A; Experimental source: strain 16M				
C; Genetics:				
A; Gene: BMET0820				
A; Map position: I				
Query Match	66.7%	Score 32;	DB 2;	Length 154;
Best Local Similarity	60.0%	Pred. No. 16;		
Matches 6;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	10		
Db 138	SLPLNADPVE	147		
RESULT 4				
Query Match	66.7%	Score 32;	DB 2;	Length 154;
Best Local Similarity	60.0%	Pred. No. 16;		
Matches 6;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	10		
Db 138	SLPLNADPVE	147		
hypothetical protein At2g16110 [imported] - <i>Arabidopsis thaliana</i>				
C; Species: <i>Arabidopsis thaliana</i> (mouse-ear cress)				
C; Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001				
C; Accession: GB4536				
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999				
A; Title: Sequence and analysis of chromosome 2 of the plant <i>Arabidopsis thaliana</i> .				
A; Reference number: AB4420; MUID:20083487; PMID:10617197				
A; Accession: G84536				
A; Status: preliminary				
A; Molecule type: DNA				
A; Residues: 1-323 <STO>				
A; Cross-references: GB:AE002093; NID:94678207; PIDN:AAD26953.1; GSPDB:GN00139				
C; Genetics:				
A; Gene: At2g16110				
A; Map position: 2				
Query Match	66.7%	Score 32;	DB 2;	Length 323;
Best Local Similarity	75.0%	Pred. No. 37;		
Matches 6;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	1		
Db 300	NPMLNVAMP	309		
RESULT 6				
Query Match	66.7%	Score 32;	DB 2;	Length 425;
Best Local Similarity	60.0%	Pred. No. 49;		
Matches 6;	Conservative 2;	Mismatches 2;	Indels 0;	Gaps 0;
Oy 1	SMLNAAV	10		
Db 300	NPMLNVAMP	309		
probable sodium- and chloride- dependent transporter - syphilis spirochete				
C; Species: <i>Treponema pallidum</i> subsp. <i>pallidum</i> (<i>Syphilis</i> spirochete)				
C; Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999				
C; Accession: CT1376				
R; Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G; Richardson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M; they, L.; Wedin, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998				
A; Title: Complete genome sequence of <i>Treponema pallidum</i> , the syphilis spirochete.				
A; Reference number: AT1250; MUID:98332770; PMID:9665876				
A; Accession: CT1376				
A; Status: preliminary; nucleic acid sequence not shown; translation not shown				
A; Molecule type: DNA				
A; Residues: 1-443 <COL>				
A; Cross-references: GB:AE001187; GB:AE000520; NID:g332273; PIDN: AAC65017.1; PID:g3322				
C; Genetics:				
A; Gene: At2g16110				
A; Map position: 2				

A;Gene: TP0023
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 66.7%; Score 32; DB 2; Length 443;
Best Local Similarity 50.0%; Pred. No. 52; Mismatches 5; Conservative 5; MisMatches 0; Indels 0; Gaps 0;

Qy 1 SMPLNAAVKM 10
Db 360 SLPNLSMRV 369

RESULT 7
B9627
hypothetical protein T8K14.1 [imported] - Arabidopsis thaliana

C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Accession: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Ruizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziali, A.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:1130712

A;Status: preliminary

A;Molecule type: DNA

A;Cross-references: GB:AE005173; NID:94835752; PIDN:AAD30219.1; GSPDB:GN00141

C;Genetics:

A;Gene: T8K14.1

A;Map position: 1

Query Match 66.7%; Score 32; DB 2; Length 1248;
Best Local Similarity 60.0%; Pred. No. 1e+02; Mismatches 6; Conservative 3; MisMatches 1; Indels 0; Gaps 0;

Qy 1 SMPLNAAVKM 10
Db 380 SLPNAYAVL 389

RESULT 8
T10673
hypothetical protein F6E21.110 - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Accession: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

R;Bevan, M.; Leonard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Banerjee, A;Submitted to the Protein Sequence Database, June 1999

A;Accession number: Z16533

A;Residues: 1-114 <BEV>

A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.110

C;Genetics:

A;Gene: ATSP:F6E21.110

A;Map position: 4

Query Match 64.5%; Score 31; DB 2; Length 114;
Best Local Similarity 75.0%; Pred. No. 20; Mismatches 6; Conservative 2; MisMatches 0; Indels 0; Gaps 0;

RESULT 9
A81899
probable amidase NMA1303 [imported] - *Neisseria meningitidis* (strain z2491 serogroup C)

C;Species: *Neisseria meningitidis*
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: A81899; E81886

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc Holland, S.; Jacobs, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandhra, N.; Rutter, S.; Simmonds, M.; Stevens, M.; Toleman, M.; van de Rijn, M.; Vaidya, A.; White, D.; Woodward, S.; Zhou, J.

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81899

A;Status: preliminary

A;Molecule type: DNA

A;Cross-references: GB:AL162755; GB:AL157059; NID:97379742; PIDN:CAB844555.1; PID:97379742

A;Residues: 1-181 <PAR>

A;Experimental source: serogroup A, strain Z2491

A;Accession: E81866

A;Status: preliminary

A;Molecule type: DNA

A;Cross-references: GB:AL162755; GB:AL157059; NID:97379742; PIDN:CAB844555.1; PID:97379742

A;Residues: 1-181 <PA2>

A;Experimental source: serogroup A, strain Z2491

A;Cross-references: GB:AL162755; GB:AL157059; NID:97379742; PIDN:CAB844555.1; PID:97379742

A;Genes: NMA1303; NMA1188

Query Match 64.6%; Score 31; DB 2; Length 181;
Best Local Similarity 60.0%; Pred. No. 33; Mismatches 6; Conservative 2; MisMatches 2; Indels 0; Gaps 0;

Qy 1 SMPLNAAVKM 10
Db 57 NMPLREAVKL 66

RESULT 10
C81813
duplicated hypothetical protein NMA1864 [imported] - *Neisseria meningitidis* (strain z2491)

C;Species: *Neisseria meningitidis*
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: C81813

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc Holland, S.; Jacobs, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandhra, N.; Rutter, S.; Simmonds, M.; Stevens, M.; Toleman, M.; van de Rijn, M.; Vaidya, A.; White, D.; Woodward, S.; Zhou, J.

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: C81813

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-181 <PAR>

A;Cross-references: GB:AL162757; GB:AL157059; NID:97380371; PIDN:CAR85087.1; PID:97380371

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1864

Query Match 64.6%; Score 31; DB 2; Length 181;
Best Local Similarity 60.0%; Pred. No. 33; Mismatches 6; Conservative 2; MisMatches 2; Indels 0; Gaps 0;

Qy 1 SMPLNAAVKM 10
Db 57 NMPLREAVKL 66

RESULT 11
T02471
hypothetical protein At2g45760 [imported] - *Arabidopsis thaliana*

N;Alternative names: hypothetical protein F4118.26

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

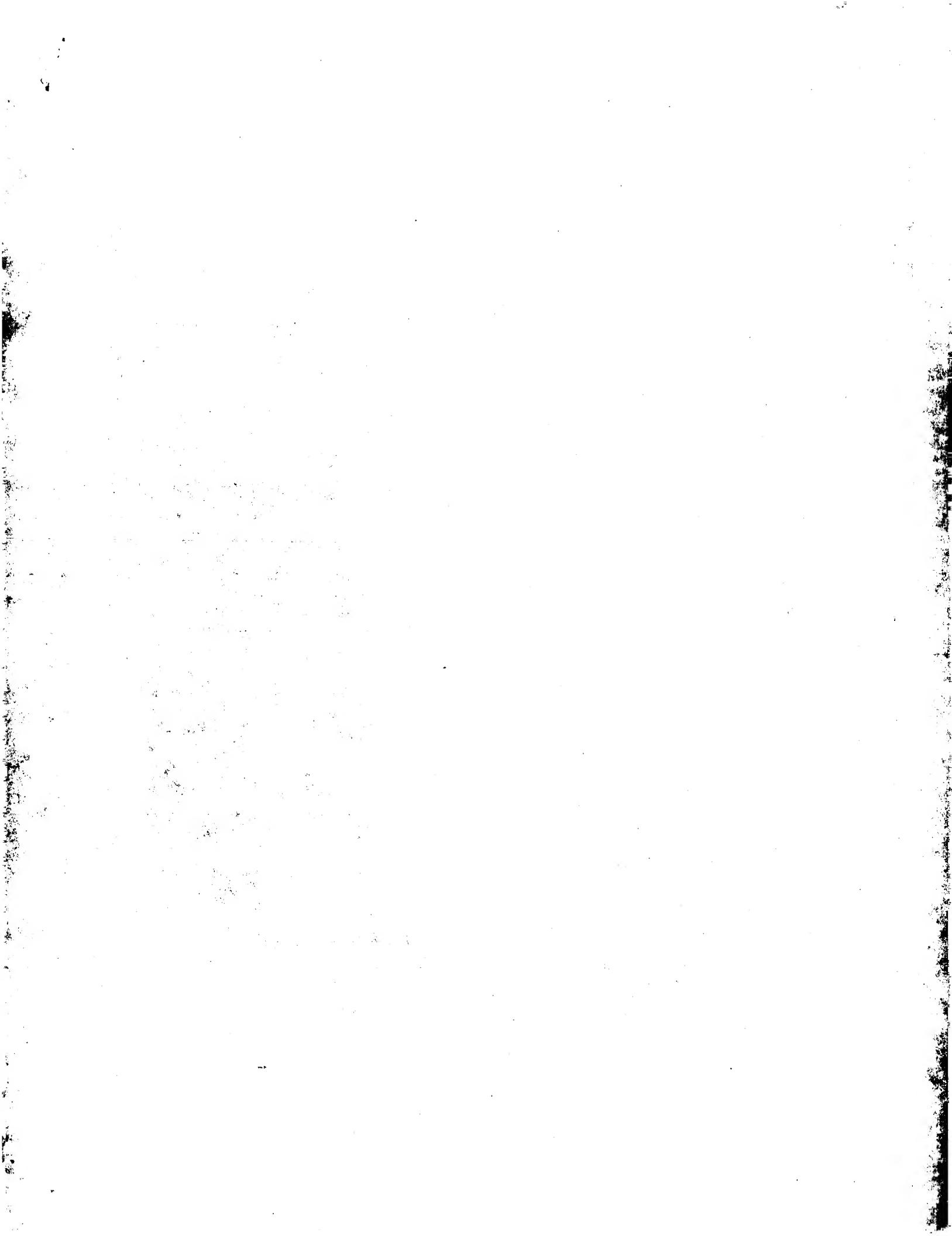
C;Accession: T02471; E84894

R;Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K

submitted to the EMBL Data Library, August 1998
A; Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.
A; Reference number: Z14574
A; Status: translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-279 <RQ>
A; Cross-references: EMBL:AC004665; NID:93386593; PID:93386618
A; Experimental source: cultivar Columbia
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shan, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.;
Nature 402, 761-768, 1999
A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A; Reference number: A84420; MUID:20083487; PMID:10617197
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-279 <STO>
A; Cross-references: GB:AE002093; NID:93386618; PIDN: AAC28548.1; GSPDB: GN00139
C; Genetics:
A; Gene: F4118_26; At2g45760
A; Map position:
A; Introns: 200/2
A; Superfamily: Arabidopsis thaliana hypothetical protein At2g45760
Query Match 64.6%; Score 31; DB 2; Length 279;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 MPLNAAVK 9
||:|||:
Db 64 MPINASVR 71
RESULT 12
S77169
hypothetical protein slr2144 - Synechocystis sp. (strain PCC 6803)
C; Species: Synechocystis sp.
A; Variety: PCC 6803
C; Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C; Accession: S77169
R; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpoo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A; Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A; Reference number: S74222; MUID:97061201; PMID:8905231
A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A; Accession: S77169
A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-301 <KAN>
A; Cross-references: EMBL:D90908; GB:AB001339; NID:91652725; PIDN:BAA17727.1; PID:9165280
C; Genetics:
A; Gene: F4118_26; At2g45760
A; Map position:
A; Introns: 200/2
A; Superfamily: Synechocystis hypothetical protein slr2144
Query Match 64.6%; Score 31; DB 2; Length 301;
Best Local Similarity 85.7%; Pred. No. 57;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SMPNNAA 7
||:||||:
Db 65 SLPNNAA 71
RESULT 13
S8496
thioredoxin reductase [NADPH] trxB [imported] - Bacillus halodurans (strain C-125)
C; Species: Bacillus halodurans
C; Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C; Accession: C84096
A; Map position: 1
R; Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A; Title: Complete genome sequence of the alkaliiphilic bacterium Bacillus halodurans a
A; Reference number: A83650; MUID:20512502; PMID:11058132
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-315 <RQ>
A; Cross-references: GB:AP001519; GB:BA000004; NID:910176109; PIDN:BAB07290.1; GSPDB:G
A; Experimental source: strain C-125
A; Genetics:
A; Gene: trxB
C; Superfamily: thioredoxin reductase; thioredoxin reductase homology
Query Match 64.6%; Score 31; DB 2; Length 315;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 MPLNAAVK 9
||:|||:
Db 243 LPNNEAVK 250
RESULT 14
E84299
hypothetical protein vng1455h [imported] - Halobacterium sp. NRC-1
C; Species: Halobacterium sp. NRC-1
C; Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C; Accession: E84299
R; Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leithauser, B.; Keller, R.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
; Jung, K.H.; Alam, M.; Freitas, T.; Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A; Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A; Title: Genome sequence of Halobacterium species NRC-1
A; Reference number: A84160; MUID:2050483; PMID:11016959
A; Accession: E84299
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-48 <STO>
A; Cross-references: GB:AE004437; NID:910580951; PIDN:AAG19761.1; GSPDB:GN00138
C; Genetics:
A; Gene: VNG1455H
Query Match 64.6%; Score 31; DB 2; Length 348;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMPNNAA 8
||:||||:
Db 33 STPLNNAI 40
RESULT 15
AF3363
membrane fusion protein mtrc [imported] - Brucella melitensis (strain 16M)
C; Species: Brucella melitensis
C; Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C; Accession: AF3363
R; DelVecchio, V.G.; Kaparad, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A; Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A; Reference number: AD3252; PMID:11756688
A; Accession: AF3363
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-364 <RUR>
A; Cross-references: GB:AE008917; PIDN:AAL52073.1; PID:917982843; GSPDB:GN00190
A; Experimental source: strain 16M
C; Genetics:
A; Gene: BMEI0892
A; Map position: 1

Query	Match	Score	DB	Length
Best Local Matches	Similarity 60.0%; Conservative 6;	31	2	364
Oy	SIMPLNAAVMK 10			
Db	: : : AMPANAVAL 135			
		2	Indels	0
		2	Mismatches	0
		0	Gap	0

Search completed: March 25, 2003, 08:22:16
 Job time : 12.901 secs



Copyright (c) 1993 - 2003	GenCore version 5.1.4_p5_4578			
	CompuGen Ltd.			
ON protein - protein search, using sw model.				
Run on:	March 25, 2003, 08:18:08 ; (without alignments)			
	Search time 6.0601 Seconds			
	68.436 Million cell updates/sec			
Title:	US-09-646-532B-1			
Perfect score:	48			
Sequence:	1 SMPLNAAVAKM 10			
Scoring table:	BLOSUM62			
	Gapop 10.0 , Gapext 0.5			
Searched:	112892 sqns, 41476328 residues			
Total number of hits satisfying chosen parameters:	112892			
Minimum DB seq length:	0			
Maximum DB seq length:	2000000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
	Listing first 45 summaries			
Database :	SWISSPROT_40;*			
Pred.	No. 1 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES				
Result No.	Score	Query Match Length	DB ID	Description
1	32	66.7	425	1 YG3Q_YEAST
2	31	64.6	877	1 SV_METH
3	31	64.6	1522	1 MDP_HABIN
4	30	62.5	181	1 Y167_UREPA
5	30	62.5	119	1 MICH_HUMAN
6	30	62.5	165	1 YF59_METHA
7	30	62.5	310	1 GH2_CORAM
8	30	62.5	423	1 YF95_YEAST
9	30	62.5	456	1 P162_SCHPO
10	30	62.5	805	1 P182_HABIN
11	30	62.5	810	1 FAF1_YEAST
12	30	62.5	830	1 ACA1_ARATH
13	30	62.5	1020	1 TEL1_MOUSE
14	30	62.5	1134	1 TEL1_BOVIN
15	30	62.5	1136	1 TEL1_HUMAN
16	30	62.5	1138	1 CIAA_SACKL
17	30	62.5	1839	1 YQ82_CABEL
18	29	60.4	128	1 ITMA_HUMAN
19	29	60.4	263	1 ITMA_MOUSE
20	29	60.4	263	1 MHD_ECOLI
21	29	60.4	269	1 USC_PSEAE
22	29	60.4	336	1 DCUP_DEARR
23	29	60.4	344	1 VLYON_BACSU
24	29	60.4	363	1 CAR1_CABEL
25	29	60.4	425	1 GNT2_HUMAN
26	29	60.4	447	1 SYC_SOJUSO
27	29	60.4	470	1 YB9_YEAST
28	29	60.4	602	1 FEOB_MEIJIA
29	29	60.4	668	1 YETL_SCHRO
30	29	60.4	693	1 SPOT_AQUAE
31	29	60.4	696	1 GSP_MCTU
32	29	60.4	941	1 MCIL_YEAST
33	29	60.4	950	1 P53258 saccharomyces
ALIGNMENTS				
RPT	YG3Q_YEAST	STANDARD;	PRT;	425 AA..
RC	ID YG3Q_YEAST			
RA	Skala J., Nawrocki A., Goffeau A.;			
RT	P39927; RT 01-FEB-1995 (Rel. 31, Created)			
RT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Hypothetical 47.0 kDa protein in CYSA - PEM1 intergenic region.			
GN	YGR156W OR G6670.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycetes;			
OC	Saccharomycetidae; Saccharomycetaceae; Saccharomyces.			
OX	NCBI TaxID:4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S28C;			
RA	Medline=96158062; PubMed=8585325;			
RA	One B.I., Inoue T., Kijima K., Matsuda A., Negishi K., Shinoda S., Submitted (JUN-1993) to the EMBL/GenBank/DDJB databases.			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	YGR156W			
RA	Skala J., Nawrocki A., Goffeau A.;			
RA	"The sequence of a 27 kb segment on the right arm of chromosome VII from Saccharomyces cerevisiae reveals MOL1, NAV2, RPL30B, RSR1, CYS4, PBM1/CH02, NSR1 genes and ten new open reading frames.";			
RL	RT Yeast 11:1421-1427(1995).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	CC			
CC	CC			
CC	CC			
DR	EMBL; D16502; BA03953; 1; -.			
DR	EMBL; X85807; CAA59813; 1; -.			
DR	EMBL; 272941; CAA97170; 1; -.			
DR	PTR; S48506; S48506.			
DR	SGD; S0003388; YGR156W			
DR	InterPro; IPR00504; RNA_rec_mot.			
DR	SMART; SM00660; RRM; 1.			
DR	KW Hypothetical protein.			
FT	CONFLICT 300 AA;			
SBQENC	N -> Y (IN RRF. 1) . CRC64;			
Query	Match 66.7%; Score 32; DB 1; Length 425;			
Matches	Best Local Similarity 60.0%; Pred. No. 20;			
6;	Conservative 2; Mismatches 2; Indels 0; Gaps 0;			
QY	1 SMPLNAAVAKM 10			
Db	300 NMPLNVAAMP 309			

RESULT 2
SYV_METH STANDARD; PRT; 877 AA.
ID SYV-METH
AC O28861;
DT 15-JUL-1998 (Rel. 36, created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE VALY-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS).
GN Methyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS).
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beta H;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubots J.,
RA Alderedge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wiezbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabakar S.,
RA Mcbougail S., Shimer G., Goyal A., Pietrovskil S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT Complete genome sequence of Methanobacterium thermoautotrophicum
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic
CC -!- SIMILARITY: BELONGS TO CLASS I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000855; AAB85270.1; -.
CC HSSP; B96142; IGAX.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR00112; tRNA-synt_I.
DR InterPro; IPR002303; tRNA-synt_val.
DR PR00133; tRNA-synt_1; 1.
DR PRINTS; PR00886; TRNASYNTHVAL.
DR TIGRFAMS; TIGR00422; valS; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 46 56 "HIGH" REGION.
FT SITE 529 533 "KMSKS" REGION.
FT BINDING 532 532 ATP (BY SIMILARITY).
SQ SEQUENCE 877 AA; 100946 MW; B939720D608FF0DB CRC64;

Query Match 64.6%; Score 31; DB 1; Length 877;
Best Local Similarity 75.0%; Pred No 72;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DY 2 MPLNAAVK 9
QY 11111:1
Db 742 MPLNAPLK 749

RESULT 3
MRP3_RAT STANDARD; PRT; 1522 AA.
ID MRP3_RAT
AC O8B563; 088270;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Canalicular multispecific organic anion transporter 2 (Multidrug

DE resistance-associated protein 3) (MRP-like protein-2) (MRP-2).
GN ABCC3 OR CM0A12 OR MRP3 OR MLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA ORLIZ D.F., Li S., Iyer R., Zhang X., Novikoff P., Arias I.M.;
RA "MRP2, a new AMP-binding cassette protein localized to the canalicular
RT domain of the hepatocyte";
RL Am. J. Physiol. 276:G1493-G1500(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Strige-Dawley; TISSUE=Colon;
RA Hirschbiel T., Suzuki H., Ito K., Ogawa K., Kume K., Shimizu T.,
RA Sugiyama Y.; Pubmed=9614210;
RT "Hepatic expression of multidrug resistance-associated protein-like
RT proteins maintained in esai hyperbilirubinemic rats";
RL Mol. Pharmacol. 53:1068-1075(1998).
CC -!- FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND
CC INTESTINAL EXCRETION OF ORGANIC ANTONS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: LUNG AND INTESTINE, LOW IN LIVER. HIGHER IN
CC LIVER OF ESAI HYPERBILIRUBINEMIC RATS AND TR(-) MUTANT RATS.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF072816; AAC25416.1; -.
DR HSSP; P13569; INBD.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR001467; BAN28955.1; -.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR00110; ABCtransportr.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF0005; ABC_tran; 2.
DR Pfam; PF00654; ABC_membrane; 2.
DR Prodom; PD000006; ABC_transportr; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMS; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00211; ABC_TRANSPORTR; 2.
DR ATP-binding; Glycoprotein; Transport; Repeat.
FT DOMAIN 1 32 TRANSMEM
FT DOMAIN 33 53 1 (BY SIMILARITY).
FT DOMAIN 54 73 CYTOPLASMIC (BY SIMILARITY).
FT DOMAIN 74 94 2 (BY SIMILARITY).
FT DOMAIN 95 99 EXTRACELLULAR (BY SIMILARITY).
FT DOMAIN 100 120 3 (BY SIMILARITY).
FT DOMAIN 121 132 CYTOPLASMIC (BY SIMILARITY).
FT DOMAIN 133 153 4 (BY SIMILARITY).
FT DOMAIN 154 171 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 172 192 5 (BY SIMILARITY).
FT DOMAIN 193 301 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 302 322 6 (BY SIMILARITY).
FT DOMAIN 323 347 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 348 368 7 (BY SIMILARITY).
FT DOMAIN 369 424 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 425 445 8 (BY SIMILARITY).
FT DOMAIN 446 448 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 449 469 9 (BY SIMILARITY).
FT DOMAIN 470 531 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 532 552 10 (BY SIMILARITY).
FT DOMAIN 553 555 EXTRACELLULAR (BY SIMILARITY).

FT TRANSMEM 575 595
 FT DOMAIN 596 958
 FT TRANSMEM 959 979
 FT DOMAIN 980 1016
 FT TRANSMEM 1017 1037
 FT DOMAIN 1038 1080
 FT TRANSMEM 1081 1101
 FT DOMAIN 1102 1102
 FT TRANSMEM 1103 1123
 FT DOMAIN 1124 1194
 FT TRANSMEM 1195 1215
 FT DOMAIN 1216 1238
 FT TRANSMEM 1218 1232
 FT DOMAIN 1239 1522
 NP BIND 659 666
 NP BIND 1318 1325
 FT CARBOHYD 18 18
 FT CARBOHYD 1001 1002
 FT CARBOHYD 1002 1002
 CONFLICT 323 344
 FT CONFLICT 645 645
 FT CONFLICT 1075 1075
 SQ SEQUENCE 1522 AA; 158977 MW;

Query Match 64.6%; Score 31; DB 1; Length 1522;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPLNNAVKM 10
 :||| |||
 Db 461 IPLNGAVAVM 469

RESULT 4
 MOAD_HAEIN STANDARD; PRT; 81 AA.

ID MOAD_HAEIN STANDARD; PRT; 81 AA.
 AC P45309;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Molybdopterin converting factor subunit 1 (MPT synthase subunit 1)
 DE (Molybdopterin synthase subunit 1) (Molybdenum cofactor biosynthesis
 protein D) (Molybdopterin converting factor small subunit).
 MOAD OR H1674.

OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TAXID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=752800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., FitzHugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Heidelberg J., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 Rd";
 RT Science 269:496-512(1995).
 RN [2]
 RP REVISIONS.

RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.;
 RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Converts molybdopterin precursor Z into molybdopterin.
 CC This requires the incorporation of two sulfur atoms into precursor
 CC Z to generate a dithiolene group. The sulfur atoms are provided by
 the active form of the small subunit (By similarity).

FT TRANSMEM 11 (BY SIMILARITY); CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 12 (BY SIMILARITY); EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 13 (BY SIMILARITY); CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 14 (BY SIMILARITY); EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 15 (BY SIMILARITY); CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 16 (BY SIMILARITY); EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 17 (BY SIMILARITY); CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 18 (BY SIMILARITY); ATP (POTENTIAL).
 FT TRANSMEM 19 (BY SIMILARITY); N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT TRANSMEM 20 (BY SIMILARITY); N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT TRANSMEM 21 (BY SIMILARITY); SPSSHTHCSASSSGGFRPPGPy -> LSFINPOLISILIRF
 FT TRANSMEM 22 (BY SIMILARITY); ISDPFAPT (IN REF. 2).
 FT TRANSMEM 23 (BY SIMILARITY); I -> L (IN REF. 2).
 FT TRANSMEM 24 (BY SIMILARITY); H -> D (IN REF. 2).
 SQ SEQUENCE 1522 AA; 158977 MW;

Query Match 64.6%; Score 31; DB 1; Length 1522;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPLNNAVK 9
 :||| :|||
 Db 60 MPLESAVK 67

RESULT 5
 Y167_UREPA STANDARD; PRT; 119 AA.

ID Y167_UREPA STANDARD; PRT; 119 AA.
 AC Q9PQX6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Y167.
 DE UU167.
 OS ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
 OC NCBITAXID=134821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serovar 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RA "The complete sequence of the mucosal pathogen Ureaplasma
 RT urealyticum.";
 RL Nature 407:757-762(2000).
 CC ---

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AE002115; AAC30574_1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 119 AA; 14280 MW;

Query Match 62.5%; Score 30; DB 1; Length 119;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MPLNNAVK 9
 :||| :|||
 Db 80 MPLNLAVO 87

RESULT 6

MICH ROMAN	STANDARD;	PRT;	165 AA.
ID P20382; Q16044;			
AC 01-FEB-1991 (Rel. 17, Created)			
DT 15-JUL-1999 (Rel. 38, Last sequence update)			
DR 15-JUN-2002 (Rel. 41, Last annotation update)			
DE PRO-MCH precursor [Contains: Neuropeptide-G-E; Neuropeptide-glycine-D glutamic acid-isoleucine (NEI); (Neuropeptide E-1); Melanin-concentrating hormone (MCH)].			
GN PMCH OR MCH.			
OC Homo sapiens (Human).			
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Cetartiodactyla; Hominoidea; Homo.			
CC NCBIX TAXID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE-HYPOTHALAMUS;			
RT MEDLINE=91125371; PubMed=2149166;			
RX Presse F., Nahon J.-L., Fischer W.H., Vale W.;			
RA "Structure of the human melanin-concentrating hormone mRNA.";			
RL Mol. Endocrinol. 4:632-637(1990).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC TISSUE-Breast cancer;			
RT MEDLINE=93316802; PubMed=8326825;			
RA Bretton C., Schorp M., Nahon J.-L.;			
RT "Isolation and characterization of the human melanin-concentrating hormone gene and a variant gene.";			
RT Brain Res. Mol. Brain Res. 18:297-310(1993).			
RN [3]			
RP PROCESSING.			
RX MEDLINE=99156937; PubMed=10037747;			
RA Viale A., Ortola C., Hervieu G., Furuta M., Barbero P., Steiner D.F., Seidan N.G., Nahon J.-L.;			
RT "Cellular localization and role of prohormone convertases in the processing of pro-melanin-concentrating hormone in mammals.";			
RT J. Biol. Chem. 274:6556-6545(1999).			
RN [4]			
RP TISSUE SPECIFICITY.			
RX MEDLINE=97334402; PubMed=9191099;			
RA Viale A., Zhiying Y., Breton C., Pedeutour F., Coquerel A., Jordan D., Nahon J.-L.;			
RT "The melanin-concentrating hormone gene in human: flanking region analysis, fine chromosome mapping, and tissue-specific expression.";			
RT Brain Res. Mol. Brain Res. 46:243-255(1997).			
CC -I- FUNCTION: MCH may act as a neurotransmitter or neuromodulator in a broad array of neuronal functions directed toward the regulation of goal-directed behavior, such as food intake, and general arousal. May also have a role in spermatocyte differentiation.			
CC -I- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LATERAL HYPOTHALAMUS, ALSO DETECTED IN PALLIDIUM, NEOCORTEX AND CEREBELLUM. ALSO FOUND IN THYMUS, BROWN ADIPOSE TISSUE, DUODENUM, AND TESTIS (SPERMATOGONIA, EARLY SPERMATOCTES AND SERTOLI CELLS). NO EXPRESSION IN PERIPHERAL BLOOD. IN BRAIN EXCLUSIVELY MATURE MCH AND NEI PEPTIDES ARE PRESENT. IN PERIPHERAL TISSUES A LARGE PRODUCT, ENCOMPASSING THE NEI AND MCH DOMAINS OF THE PRECURSOR, IS FOUND PREDOMINANTLY.			
CC -I- TIM: DIFFERENTIALLY PROCESSED IN THE BRAIN AND IN PERIPHERAL ORGANS PRODUCING TWO NEUROPEPTIDES; NEI AND MCH. A THIRD PEPTIDE, NEGE, MAY ALSO BE PRODUCED. PREFERENTIAL PROCESSING IN NEURONS BY PROHORMONE CONVERTASE 2 (PC2) GENERATES NEI. MCH IS GENERATED IN CONVERTASES INCLUDING PC1/3, PC2 AND PC5/6.			
CC -I- SIMILARITY: BELONGS TO THE MELANIN-CONCENTRATING HORMONE FAMILY.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC CC DR EMBL; M57703; AA63214.1; -.			
CC CC DR EMBL; S63697; AAB27493.1; -.			
CC CC DR PTB; A34551; A34551.			
CC CC DR Genew; HGNC:9109; PMCH.			
CC CC DR MM: 176795; -.			
CC KW Cleavage on pair of basic residues; Hormone; Neuropeptide; Signal; KW SIGNAL; Spermatogenesis.			
CC KW POTENTIAL.			
CC FT SIGNAL 1 21			
CC FT PEPTIDE 22 165			
CC FT PEPTIDE 110 128			
CC FT PEPTIDE 131 143			
CC FT PEPTIDE 147 165			
CC FT MOD_RES 143 143			
CC FT DISULFID 153 162			
CC FT CONFLICT 42 42			
CC FT CONFLICT 104 106			
CC FT CONFLICT 113 113			
CC FT SEQUENCE 165 AA; 18723 MW; D639289386637244 CRC64;			
CC SO Query Match 62.5%; Score 30; DB 1; Length 165;			
CC SO Best Local Similarity 62.5%; Pred. No. 20; Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			
CC OQ 2 MPINAAYK 9			
CC Db 96 iPLNLAIK 103			
CC RESULT /			
CC YP85_METJA			
CC ID YP85_METJA STANDARD; PRT; 310 AA.			
CC AC Q58980;			
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)			
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)			
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)			
CC DE Putative aldolase MJ1585 (EC 4.2.1.-).			
CC GN MJ1585.			
CC OS Methanococcus jannaschii.			
CC OC Archaea; Euryarchaeota; Methanococci; Methanococcales;			
CC OC Methanococcaceae; Methanococcaceae; Methanococcaceae; NCBIX TAXID=2190;			
CC RN [1]			
CC RP SEQUENCE FROM N.A.			
CC RC STRAIN-JALL-1 / DSM 2661 / ATCC 43067;			
CC RX MEDLINE=96337999; PubMed=1688037;			
CC RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D., Utterback T.R., Kelley J.M., Peterson D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kline B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Wosse C.R., Venter J.C., RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.", Science 273:1058-1073(1996).			
CC RT Jannaschii			
CC RL Science 273:1058-1073(1996).			
CC CC -I- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.			
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC CC DR EMBL; U67598; AAB99604.1; -.			
CC CC DR TrIGR; MJ1585; -.			
CC CC DR InterPro; IPR002915; DeoC.			
CC CC DR Pfam; PF01791; DeoC; 1.			

KW HYPOTHETICAL PROTEIN; Lyase; Schiff base; Complete proteome.
 FT BINDING_213 213 SCHIFF-BASE (BY SIMILARITY).
 SQ SEQUENCE 310 AA; 34574 MW; 4AB6286F888049c0 CRC64;

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Matches
Oy 3	PNAAVKM 10	62.5%	30	1	30	39	5; Conservative
Db 4	PLNSAVRL 11	62.5%	55	3	60	60	3; Mismatches

RESULT 8
 GCH2_CORAM STANDARD; PRT; 423 AA.
 ID GCH2_CORAM AC 024752; DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DR 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Riboflavin biosynthesis protein ribA [Includes: GTP cyclohydrolase II
 (EC 3.5.4.25); 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBP
 synthase)].
 DE RIBA.
 OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
 OC Bacteria; Actinobacteria; Actinomycetales; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TAXID=1697;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=S288C / As97;
 RX MEDLINE-9733271; Published=9169875;
 RA Bussey R., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Barnes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPolo T., Dubois E., Dussterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffau A.,
 RA Hall J., Hebling U., Heumann K., Hibert B., Hillier L.,
 RA Hunicke-Smith S., Hyman R., Jobsthan M., Kalman S., Kleine K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Matthe R., Messenguy F., Mewes H.-W., Mittipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Schafe M.,
 RA Schierens B., Schramm S., Schroeder M., Sdcu A.M., Tettelin H.,
 RA Urestarazu R., Ushinsky S., Viereckels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Weiler H., Winnett E.,
 RA Zhang W.W., Zolnier A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 CC -1- SIMILARITY: BELONGS TO THE DHBP
 CC -1- PATHWAY: Riboflavin biosynthesis, 31-DEC-1996
 CC -1- CATALYTIC ACTIVITY: GTP + 3-H(2O) = formate + 2,5-diamino-6-
 CC hydroxy-4-(5'-phosphorylosyl)aminopyrimidine + diphosphate.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GTP
 CC -1- CYCLOHYDROLASE II FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB003693; BAA20055.; -
 DR InterPro; IPR000422; DHBP_synthase.
 DR Pfam; PF00925; GTP_cyclohydro2; 1.
 DR Pfam; PF00926; DHBP_synthase; 1.
 DR ProDom; PD003034; DHBP_synthase; 1.
 DR InterPro; IPR000926; GTP_cyclohydro2.
 DR Pfam; PF00925; GTP_cyclohydro2; 1.
 DR Pfam; PF00926; DHBP_synthase; 1.
 DR TIGRFAMS; TIGR00505; ribA; 1.
 DR TIGRFAMS; TIGR00506; ribB; 1.
 KW Multifunctional enzyme; Riboflavin biosynthesis; Hydrolase.
 FT DOMAIN 1 210 DOMAIN 211 423 DOMAIN 212 46399 MW; EBEB9CA3E902BD25 CRC64;
 FT SEQUENCE 423 AA; 46399 MW; EBEB9CA3E902BD25 CRC64;

RESULT 10
 Query Match 62.5%; Score 30; DB 1; Length 423;
 Best Local Similarity 62.5%; Pred. No. 55; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Matches
Oy 3	PNAAVKM 10	62.5%	30	1	30	39	5; Conservative
Db 4	PLNSAVRL 11	62.5%	55	3	60	60	3; Mismatches

RESULT 10
 Query Match 62.5%; Score 30; DB 1; Length 456;
 Best Local Similarity 75.0%; Pred. No. 60; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Matches
Oy 3	PNAAVKM 10	62.5%	30	1	30	39	5; Conservative
Db 4	PLNSAVRL 11	62.5%	55	3	60	60	3; Mismatches

YPI095C OR LPG13C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TAXID=9932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=S288C / As97;
 RX MEDLINE-9733271; Published=9169875;
 RA Bussey R., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Barnes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPolo T., Dubois E., Dussterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffau A.,
 RA Hall J., Hebling U., Heumann K., Hibert B., Hillier L.,
 RA Hunicke-Smith S., Hyman R., Jobsthan M., Kalman S., Kleine K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Matthe R., Messenguy F., Mewes H.-W., Mittipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Schafe M.,
 RA Schierens B., Schramm S., Schroeder M., Sdcu A.M., Tettelin H.,
 RA Urestarazu R., Ushinsky S., Viereckels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Weiler H., Winnett E.,
 RA Zhang W.W., Zolnier A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 CC -1- SIMILARITY: BELONGS TO THE UFF0017 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U43281; AAB68204; 1; -
 DR SGD; S0005016; YPI095C.
 DR InterPro; IPR00073; Abhydrolase.
 DR InterPro; IPR000379; Ser_Esters_site.
 DR InterPro; IPR000552; UPR0017.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS01133; UPF0017; 1.
 DR InterPro; IPR000379; Ser_Esters_site.
 DR InterPro; IPR000552; UPR0017.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS01133; UPF0017; 1.
 SQ SEQUENCE 456 AA; 51723 MW; CABC660FF44A287E CRC64;

RESULT 10
 Query Match 62.5%; Score 30; DB 1; Length 456;
 Best Local Similarity 75.0%; Pred. No. 60; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Query Match	Best Local Similarity	Score	DB	Length	Pred.	No.	Matches
Oy 3	PNAAVKM 10	62.5%	30	1	30	39	5; Conservative
Db 4	PLNSAVRL 11	62.5%	55	3	60	60	3; Mismatches

YPI095C OR LPG13C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TAXID=9932;

RESULT 9

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX Schizosaccharomyces;
 NCBI_TaxID=4696;

[1] SEQUENCE FROM N.A.

RN RP Zhou J.-O., Qi H., Schulz V.P., Mateyak M.K., Monson E.K.,
 RC RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., Huckle E.J., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy I., Niblett D., McDonald S., McLean J.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinowitsch E.,
 RA Rutherford K., Rutters S., Saunders D., Seeger K., Sharp S.,
 RA Skeletton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volcker G., Hart R., Robben J., Grymonpre B.,
 RA Weltjens I., Vansteensel E., Rieger M., Scheffer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Bozym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardon C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., Mccombe W.R., Paulein T., Potashkin J.,
 RA "The genome sequence of the Schizosaccharomyces pombe.";
 RT Nature 415:871-880(2002).

CC !-! FUNCTION: Required for both repair of mitochondrial DNA and
 CC recognition of a recombinogenic signal characterized by a 20-bp
 CC palindromic at sequence in the ery region of mitochondrial DNA.
 CC This is a single stranded DNA-dependent ATPase and DNA helicase
 CC used by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF074944; AAC26139.1; -.

DR EMBL: AJ003388; CAZ2189.1; -.

KW Mitochondrion; ATP-binding; DNA-binding; DNA repair;
 KW DNA recombination; Helicase; Transit peptide.

FT TRANSIT ? 805 MITOCHONDRIAL (POTENTIAL).
 FT CHAIN ? 805 DNA REPAIR AND RECOMBINATION PROTEIN
 FT PIFL.

FT NP_BIND 332 339 ATP (PROBABLE).

FT DNA_BIND 744 763 POTENTIAL.

FT CONFLICT 203 203 T->A (IN REF. 1).

FT SEQUENCE 805 AA: 90049 MW: AB6DA44B47BCABC CRC64;

Query Match 62.5%; Score 30; DB 1; Length 810;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 12

FAR1_YEAST ID FAR1_YEAST STANDARD: PRT: 830 AA.

AC P21268;

DT 01-NAY-1991 (Rel. 18, created)

DT 01-NOV-1995 (Rel. 32, last sequence update)

DT 16-OCT-2001 (Rel. 40, last annotation update)

QY 1 SMP1NAAVK 9 :1:1:1|

Db 279 SVPLNSASK 287

RESULT 11

PLSB_HAEIN ID PLSB_HAEIN STANDARD: PRT: 810 AA.

AC P44837;

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glyceral-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).

GN PLSB OR HIT048;

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=Rd / KW0 / ATCC 51907;

RC MEDLINE=95350650; PUBMED=542800;

RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Weidman J.F., Phillips C.A., Spriggs T., Hebbelom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geochagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Ventur J.C.; RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd"; RT Science 269:496-512(1995).

RL CC !-! CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-acyl-sn-glycerol 3-phosphate.

CC CC !-! PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.

CC !-! SUBCELLULAR LOCATION: Membrane-bound (By similarity).

CC !-! SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U32758; AAC22406.1; -.

DR TIGR_HI0748; InterPro: IPR002123; Acyltransferase.

DR Pfam: PF01553; Acyltransferase; Lipid transferases; Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane; Complete proteome.

SQ SEQUENCE 810 AA: 92774 MW: DB8564BC3E2C370D CRC64;

DIE Cyclin-dependent kinase inhibitor FARI (CKI FARI) (Factor arrest protein).

DE FARI OR YJL157C OR J0565.

GN Saccharomyces cerevisiae (Baker's yeast).

OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetidae; Saccharomycetaceae; Saccharomyces.

OC NCBI_TaxID=4932;

RN [1] SEQUENCE FROM N.A.

RN MEDLINE#=9107938; PubMed=2147873;

RA Chang F.; Herskowitz I.;

RT "Identification of a gene necessary for cell cycle arrest by a negative growth factor of yeast: FARI is an inhibitor of a G1 cyclin, CLN2." Cell 63:999-1011(1990).

RN [2] REVISIONS TO N-TERMINUS.

RX MEDLINE#=95257932; PubMed=7739534;

RA McKinney J.D.; Cross F.R.;

RT "FARI and the G1 phase specificity of cell cycle arrest by mating factor in *Saccharomyces cerevisiae*." Mol. Cell. Biol. 15:2509-2516(1995).

RN [3] SEQUENCE FROM N.A.

RA Obermaier B.; Piravandi E.; Riske M.; Domdey H.;

RN Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.

RN [4] REVIEW.

RN Chang F.; Ra R.;

RT "Stop that cell cycle." R. Currr. Biol. 3:693-695(1993).

RN [5] CHARACTERIZATION.

RX MEDLINE#=9445395; PubMed=8066461;

RA Peter M.; Herskowitz I.; Peter M.;

RT "Direct inhibition of the yeast cyclin-dependent kinase Cdc28-Cln by Fari." R. Science 265:1228-1231(1994).

RN [6] CHARACTERIZATION.

RX MEDLINE#=96074755; PubMed=7490290;

RA Valtz N.; Peter M.; Herskowitz I.;

RT "FARI is required for oriented polarization of yeast cells in response to mating pheromones." R. J. Cell Biol. 131:863-873(1995).

RN [7] CHARACTERIZATION.

RX MEDLINE#=97994705; PubMed=9251033;

RA Valtz N.; Peter M.;

RT "Functional analysis of FARI in yeast." R. Meth. Enzymol. 283:350-365(1997).

RN [8] CHARACTERIZATION.

RX MEDLINE#=98037745; PubMed=9367996;

RA Henchoz S.; Chi Y.; Catarin B.; Herskowitz I.; Deshaies R.J.;

RT Peter M.;

RT "Phosphorylation- and ubiquitin-dependent degradation of the cyclin-dependent kinase inhibitor Far1p in budding yeast." R. Genes Dev. 11:3046-3060(1997).

RN [9] MUTAGENESIS OF SER-87 AND THR-306.

RX MEDLINE#=98298210; PubMed=9632750;

RA Gartner A.; Jovanovic A.; Jeoung D.I.; Bourlat S.; Cross F.R.;

RA Numerer G.;

RT "Pheromone-dependent G1 cell cycle arrest requires Far1 phosphorylation, but may not involve inhibition of Cdc28-Cln2 kinase, in vivo." R. Mol. Cell. Biol. 18:3681-3691(1998).

CC -!- FUNCTION: INHIBITOR OF THE CYCLIN-DEPENDENT KINASE CDC28.

CC NECESSARY FOR CELL CYCLE ARREST. INVOLVED IN PHEROMONE RESPONSE.

CC CONTRIBUTES TO MATING EFFICIENCY REQUIRED FOR ORIENTED POLARIZATION OF YEAST CELLS IN RESPONSE TO MATING PHEROMONES. SOME FARI MUTANTS APPEAR TO BE DEFECTIVE IN MATING BECAUSE THEY ARE

CC UNABLE TO LOCATE THE MATING PARTNER.

CC -!- SUBUNIT: ASSOCIATES WITH THE CDC28-CLN COMPLEX.

CC -!- INDUCTION: BY ALPHA-FACTOR IN A CELLS.

CC -!- DOMAIN: THERE IS EVIDENCE TO SUGGEST THAT THE N-TERMINAL MAY BE SUFFICIENT FOR CELL CYCLE ARREST AND THE C-TERMINAL MAY BE NECESSARY FOR SOME STEP IN MATING.

CC PTM: THOUGHT TO BE PHOSPHORYLATED BY MAP KINASE FUS3. THOUGHT TO ENHANCE THE BINDING OF FARI TO G1-SPECIFIC CYCLIN-DEPENDENT KINASE (CDK) COMPLEXES.

CC -!- SIMILARITY: TO YEAST SIE5.

CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ist-sib.ch/announce/> or send an email to license@ist-sib.ch).

CC DR SMART; SMD0184; RING_1.

CC DR PROSITE; PS0083; ZF_RING_2; 1.

CC DR Protein kinase inhibitor; Cell cycle; Cell division; Phosphorylation; Zinc-finger.

FT ZNFING 202 252 RING-TYPE.

FT MOD_RES 87 87 PHOSPHORYLATION (BY CDC28) (PROBABLE).

FT MOD_RES 306 306 S -> P (IN FAR1-22P; INDUCES CELL CYCLE ARREST IN THE ABSENCE OF ALPHA-FACTOR).

FT VARIANT 87 87 S -> A; PREVENTS CELL CYCLE-DEPENDENT DEGRADATION OF FARI.

FT CONFLICT 20 20 T->A: ABOLISHES G1 ARREST FUNCTION.

FT CONFLICT 568 568 D -> N (IN REF. 1).

FT P -> R (IN REF. 1).

SQ SEQUENCE 830 AA: 94572 MN; 834EBB2D7BB9641D CRC64;

Query Match 62.5% Score 30; DB 1; Length 830;

Best Local Similarity 71.4%; Pred. No. 1.1e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLNAVK 9

Db 751 PLNASIK 757

RESULT 13

ACAL_ARATH

ID ACAL_ARATH STANDARD; PRM; 1020 AA.

AC Q37145; Q37146; Q42571; Q9SFY1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Calcium-transferring ATPase 1, plasma membrane-type (EC 3.6.3.8)

DE (Ca²⁺-ATPase isoform 1) (Plastid envelope ATPase 1).

DE GN ACAL OR PEAL OR ATIG2770 OR T22C5.23 OR F28L5.1.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC OX NCBI_TaxID=3702;

RN [1] PRELIMINARY SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=CV. Columbia;

RX MEDLINE#=94052104; PubMed=8234257;

RA Huang L.; Barkelman T.; Franklin A.E.; Hoffman N.E.;

RT "Characterization of a gene encoding a Ca(2+)-ATPase-like protein in the plastid envelope." R.

RL PROC. NATL. ACAD. SCI. U.S.A. 90:10066-10070(1993).
 RN [2]
 RP REVISIONS.
 RA Huang L.; Berkelman T.; Franklin A.E.; Hoffman N.E.;
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9664-9664(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=21/016719; PubMed=11130712;
 RA Theologis A.; Ecker J.R.; Palm C.J.; Fedderspiel N.A.; Kaul S.;
 White O.; Alonso J.J.; Altafai H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
 Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.F.; Chin C.W.;
 Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Cleasy T.H.; Dewar K.;
 Dunn P.; Egwu P.; Feldblum T.V.; Feng J.-D.; Fong B.; Fujii C.Y.;
 Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Huijar L.;
 Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
 Kim C.J.; Koo H.L.; Kremenskaya I.; Kurtz D.B.; Kwan A.; Lam B.;
 Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
 Lin X.; Liu S.X.; Liu Z.A.; Lueros J.S.; Maiti R.; Marzali A.;
 Millitscher J.; Miranda M.; Nguyen M.; Nieman W.C.; Osborne B.I.;
 Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
 Sakano H.; Salzberg S.L.; Schwartz J.R.; Shin P.; Southwick A.M.;
 RA Sun H.; Tallon L.J.; Tambung J.; Toriumi M.; Town C.D.; Utterback T.; Van Aken S.; Vaysberg M.; Vysotskaya V.S.; Walker M.;
 RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;
 RT "Sequence analysis and analysis of chromosome 1 of the plant *Arabidopsis thaliana*";
 RL Nature 403:816-820(2000).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CITOSOL OUT OF THE CELL OR INTO ORGANELLES.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+)(In) = ADP + phosphate + Ca(2+)(Out).
 CC -!- ENZYME REGULATION: ACTIVATED BY CALMODULIN (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CHLOROPLAST ENVELOPE (INNER MEMBRANE) (PROBABLE).
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVELS IN ROOTS THAN IN LEAVES.
 CC -!- DOMAIN: THE N-terminus CONTAINS AN AUTOINHIBITORY CALMODULIN-BINDING DOMAIN, WHICH BINDS CALMODULIN IN A CALCIUM-DEPENDENT FASHION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (El-E2 ATPases). SUBAMILY IIB.
 CC CAUTION: Ref.3 sequence differs from that shown due to erroneous gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isdb-sib.ch/announce/> or send an email to license@isdb-sib.ch).
 CC -----
 DR EMBL; LO8468; RAD10211.1; -.
 DR EMBL; LO8469; RAD10212.1; -.
 DR EMBL; D11983; BAA0090.1; -.
 DR EMBL; D11984; BRA0091.1; -.
 DR EMBL; X69340; CAA449558.1; -.
 DR EMBL; X69341; CAA449559.1; -.
 DR EMBL; X69345; AAC24958.1; -.
 DR EMBL; AC019280; AAC50579.1; -.
 DR HSSP; P01491; 1EUL
 DR InterPro; IPR001757; ATPase_El-E2.
 DR InterPro; IPR004014; Cation_ATPase.
 DR InterPro; IPR001454; Hydrolase_hydriase.
 DR Pfam; PF00122; El-E2_ATPase; 1.
 DR Pfam; PF00689; Cation_ATPase; C; 1.
 DR Pfam; PF00690; Cation_ATPase_N; 1.
 DR PRINTS; PR00119; CAAATPASE.
 DR TIGRFAMS; TIGR0115; Ca_ATPase; 1.
 DR PROSITE; PS00154; ATPASE_El-E2; 1.
 DR PROSITE; PS00154; ATPASE_El-E2; 1.

 RESULT 14
 TIEL_MOUSE ID TIEL_MOUSE STANDARD; PRT; 1134 AA.
 AC Q06806; DT 01-FEB-1995 (Rel. 31, Created);
 OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI-TAXID=10090;
 GN TIE1 OR TIE-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI-TAXID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Lung;
 RX MEDLINE=94022314; PubMed=8415706;
 RA Sato T.N.; Qin Y.; Kozak C.A.; Andrus K.L.;
 RT "tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes expressed in early embryonic vascular system";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Bone marrow;
 RA Kriakov A.V.; Brshner M.A.; Visser J.W.M.; Belyavsky A.V.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

SEQUENCE OF 1-19 FROM N.A.

STRAIN-129/Sv; TISSUE-Liver;

MEDLINE-95383653; PUBMED-7655012;

Korhonen J., Lahtinen I., Halmekyto M., Alhonen L., Janne J., Dumont D., Alitalo K.;

Dumont D., Alitalo K.;

RT Endothelial-specific gene expression directed by the tie gene promoter *in vivo*.;

RL BLOOD 86:1828-1835(1995).

-I- FUNCTION: PROBABLE PROTEIN TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.

-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

-I- SUBFAMILY:

-I- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR ENDOTHELIAL CELLS.

-I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TIE SUBFAMILI.

-I- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

-I- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

SEQUENCE FROM N.A.

CC CAA50556.1; -.

CC X80764; CAA56739.1; -.

CC EMBL-X73960; CAA52148.1; -.

CC S79346; ; NOT_ANNOTATED_CDS.

CC MGD: MGI:99906; Tie1.

DR InterPro; IPR00561; EGF-like.

DR InterPro; IPR00719; Euk_pk kinase.

DR InterPro; IPR03961; FN_III.

DR InterPro; IPR03599; Ig.

DR InterPro; IPR03006; Ig_MHC.

DR InterPro; IPR03600; Ig_Like.

DR InterPro; IPR01245; Tyr_pk kinase.

PFam: PF00008; EGF; 2.

PFam: PF00041; fn3; 3.

PFam: PF00047; Ig; 2.

PFam: PF00069; pk kinase; 1.

PRINS; PR0010; TYRKINASE.

DR Prodrom; P00001; Euk_pk kinase; 1.

SMART; SMO0181; EGF; 2.

SMART; SMO0050; FN3; 2.

SMART; SMO0409; Ig; 1.

SMART; SMO0410; Ig_Like; 1.

SMART; SMO0219; TYRK; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00109; PROTEIN_KINASE_TIR; 1.

PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

PROSITE; PS00022; EGF; 1; 3.

DR PROSTE; PS01486; EGF; 2; 3.

KW Receptor; Tyrosine-protein kinase; Transferase; Signal; ATP-binding; Glycoprotein; Phosphorylation; Multigene family.

FT SIGNAL; 1 22

FT CHAIN; 23 1134 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.

FT DOMAIN; 23 755 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 756 780 POTENTIAL.

SEQUENCE OF 1-19 FROM N.A.

STRAIN-129/Sv; TISSUE-Liver;

MEDLINE-95383653; PUBMED-7655012;

Korhonen J., Lahtinen I., Halmekyto M., Alhonen L., Janne J., Dumont D., Alitalo K.;

Dumont D., Alitalo K.;

RT Endothelial-specific gene expression directed by the tie gene promoter *in vivo*.;

RL BLOOD 86:1828-1835(1995).

-I- FUNCTION: PROBABLE PROTEIN TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.

-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

-I- SUBFAMILY:

-I- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR ENDOTHELIAL CELLS.

-I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TIE SUBFAMILY.

-I- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

-I- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X71423; CAA50554.1; - .
 DR PIR; S32690; S32690.
 DR HSSP; P11362; LFGR.
 DR InterPro; IPR00561; EGF-like.
 DR InterPro; IPR00719; Euk_pk kinase.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001245; Tyr_pk kinase.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00041; fn3_3.
 DR Pfam; PF00047; Ig_2.
 DR Pfam; PF00069; Pkinase_1.
 DR PRINTS; PRO0109; TRKINASE.
 DR PRODOM; PD000001; Euk_pk kinase; 1.
 DR SMART; SM00181; EGF_2.
 DR SMART; SM00060; FN3_2.
 DR SMART; SM00409; Ig_1.
 DR SMART; SM00219; TYRKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00022; EGF_1_3.
 DR PROSITE; PS01186; EGF_2_3.
 KW Receptor; Tyrosine-protein kinase; transferase; Signal; Atm-binding;
 KW EGFR-like domain; transmembrane; Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 1136
 FT DOMAIN 24 757 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.
 FT TRANSMEM 758 782 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 43 1136 POTENTIAL.
 FT DOMAIN 43 106 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 212 254 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 256 301 EGF-LIKE 1.
 FT DOMAIN 303 343 EGF-LIKE 2.
 FT DOMAIN 370 424 EGF-LIKE 3.
 FT DOMAIN 444 538 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 541 637 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 641 742 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 837 1116 FIBRONECTIN TYPE-III 3.
 FT NP_BIND 843 851 PROTEIN KINASE.
 FT BINDING 866 866 ATP (BY SIMILARITY).
 FT ACT SITE 977 977 ATP (BY SIMILARITY).
 FT MOD_RES 1005 1005 BY SIMILARITY.
 FT CARBOHYD 84 84 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 707 707 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1136 AA; 124953 MW; EFF8504A041BB12 CRC64;

Query Match Similarity: 55.6%; Score: 30; DB: 1; Length: 1136;
 Best Local Similarity: 55.6%; Pred. No.: 1.6e-02; Indels: 0; Gaps: 0;
 Matches: 5; Conservative: 3; Mismatches: 1;

Qy	2	MPLNAVKM	10	:	
Db	861	LKMMRAIKM	869		

Search completed: March 25, 2003, 08:20:10
 Job time : 8.06061 secs

Copyright (c)1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 08:19:47 ; Search time 22.1212 Seconds

(without alignments) 93.145 Million cell updates/sec

Title: US-09-646-532B-1

Perfect score: 48

Sequence: 1 SMPLNAVKM 10

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 671580 seqs, 20647115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	RESULT 1	ALIGNMENTS
1	37	77.1	154	16	Q98MA8	Q98MA8	PRELIMINARY; ID: Q98MA8; DT: 01-OCT-2001 (TREMBLrel. 18, Created)	17	66.7 443 16 083067 treponema p
2	35	75.0	176	4	Q16051	Q16051	hom sapien	18	66.7 1104 5 0966JB
3	36	75.0	1009	4	Q14527	Q14527	hom sapien	19	66.7 1248 10 09SAJ2
4	36	75.0	1009	4	Q96KM9	Q96KM9	homo sapien	20	64.6 114 10 09M083
5	35	75.0	1009	4	Q14536	Q14536	homo sapien	21	64.6 181 16 09JR25
6	35	72.9	836	6	Q95217	Q95217	oryctolagous	22	64.6 181 16 09JR25 neisseria m
7	35	72.9	1005	6	Q95216	Q95216	oryctolagous	23	64.6 279 10 080843
8	34	70.8	450	16	Q8RDK9	Q8RDK9	thermoanaerobacter	24	64.6 301 16 073681
9	34	70.8	1134	16	Q92Rw9	Q92Rw9	rhizobium m	25	64.6 315 16 09K703
10	33	68.8	154	16	Q92057	Q92057	rhizobium m	26	64.6 348 17 09HPV6
11	33	68.8	459	16	Q99RU7	Q99RU7	staphylococ	27	64.6 417 2 09YHBO
12	33	68.8	487	13	Q90Z15	Q90Z15	xenopus lae	28	64.6 417 2 09f708
13	33	68.8	639	2	Q87538	Q87538	shewanella	29	64.6 421 16 09X296
14	33	68.8	1670	10	Q9LE42	Q9LE42	arabidopsis	30	64.6 425 16 091x03
15	32	66.7	154	16	Q8YHh9	Q8YHh9	arabidopsis	31	64.6 425 16 091x03 bacillus ha
16	32	66.7	323	10	Q9X1h8	Q9X1h8	arabidopsis	32	64.6 425 16 091x03 halobacteri
								33	64.6 425 16 091x03 brucella me
								34	64.6 425 16 091x03 uncultured
								35	64.6 503 10 09SX12
								36	64.6 506 11 09CWJ7
								37	64.6 509 11 09z79 helicobacte
								38	64.6 509 11 02542 helicobacte
								39	64.6 509 11 08x0m7 neurospora
								40	64.6 567 3 08uas3 aerobacteri
								41	64.6 650 10 09f708
								42	64.6 656 17 027609
								43	64.6 707 2 09x5r7 streptomyce
								44	64.6 716 5 09w2d2 drosophila
								45	64.6 726 3 096ur7 magnaporthe

RL	J. Biol. Chem.	270:4575-4587(1995);	the SPH motifs of the SV40 enhancer and to the HIV-1 promoter. ;
CC	-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.		
DR	EMBL: Z46606; CAA86571; 1; -;		
DR	EMBL: Z46605; CAA86572; 1; -;		
DR	TRANSFAC; T04146; -;		
DR	INTERPRO; IPR00410; DEAD.		
DR	INTERPRO; IPR001650; Helicase_C.		
ID	095216	PRELIMINARY;	PRT; 1005 AA.
AC	095216		
DT	01-FEB-1997 (TREMBLrel. 02, Created)		
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	RUSH1alpha.		
DR	PFam; PF00271; helicase_C; 1.		
DR	Pfam; PF00176; SNF2_N; 2.		
KW	ATP-binding; Helicase; Zinc-finger.		
SQ	SEQUENCE 1009 AA; 113914 MW; B4ACD86A6844CC5F CRC64;		
Query Match	75.0%; Score 36; DB 4; Length 1009;		
Best Local Similarity	70.0%; Pred. No. 45;		
Matches	7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;		
OY	1 SMPLNAVKM 10		
Db	196 SMPVHAQW 205		
RESULT 6			
ID	095217	PRELIMINARY;	PRT; 836 AA.
AC	095217;		
DR	01-FEB-1997 (TREMBLrel. 02, Created)		
DR	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	RUSH1beta.		
DE	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
NCBI TaxID=9986;	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=NEW ZEALAND WHITE;		
RX	MEDLINE=97082217; PubMed=8923460;		
RA	Hayward-Lester A., Hewettson A., Beale E.G., Oefner P.J., Doris P.A., Chilton B.S.;		
RT	"Cloning, characterization, and steroid-dependent posttranscriptional processing of RUSH-1 alpha and beta, two uteroglobin promoter-binding proteins".		
RT	Mol. Endocrinol. 10:1335-1349(1996).		
RT	SEQUENCE FROM N.A.		
RC	STRAIN=NEW ZEALAND WHITE;		
RX	MEDLINE=97082217; PubMed=8923460;		
RA	Hayward-Lester A., Hewettson A., Beale E.G., Oefner P.J., Doris P.A., Chilton B.S.;		
RT	"Cloning, characterization, and steroid-dependent posttranscriptional processing of RUSH-1 alpha and beta, two uteroglobin promoter-binding proteins".		
RT	Mol. Endocrinol. 10:1335-1349(1996).		
RL	INTERPRO: CONTAINS 1 RING-TYPE ZINC FINGER.		
DR	HSSP; P15919; IRMD.		
DR	INTERPRO; IPR00163; DEAD.		
DR	INTERPRO; IPR001410; DEAD.		
DR	INTERPRO; IPR000330; SNF2_N.		
DR	INTERPRO; IPR001841; Zinc_finger.		
DR	Pfam; PF00176; SNF2_N; 1.		
DR	Pfam; PF00097; zf-C3HC4; 1.		
DR	SMART; SM00487; DEXDC; 1.		
DR	SMART; SM00490; HELICC; 1.		
DR	SMART; SM00184; RING; 1.		
DR	PROSITE; PS00518; ZF_RING_1; 1.		
KW	ATP-binding; Helicase; Zinc-finger.		
SQ	SEQUENCE 1005 AA; 113582 MW; C741E7117D6BD807 CRC64;		
Query Match	72.9%; Score 35; DB 6; Length 1005;		
Best Local Similarity	60.0%; Pred. No. 74;		
Matches	6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;		
OY	1 SMPLNAVKM 10		
Db	196 SMPVHAQW 205		
RESULT 8			
ID	08RDK9	PRELIMINARY;	PRT; 450 AA.
AC	08RDK9;		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
GN	UDP-N-acetylglucosaminyl tripeptide synthase.		
GN	MUR OR TEP0008.		
OS	Thermanaerobacter tengcongensis.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia; Thermanaerobacteriales; Thermanaerobacteriaceae; Thermanaerobacter.		
OC	NCBI_TaxID=119072;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MBAT / JCM11007;		
RX	MEDLINE=21992816; PubMed=11997336;		
RA	Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J., Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L., Tan H., Chen R., Wang J., Yu J., Yang H., "A complete sequence of T. tengcongensis genome.";		
OY	1 SMPLNAVKM 10		

RL Genome Res. 12:689-700(2002).
 DR EMBL; AE012914; AACM23325.1; -.
 KW Complete proteome.
 SQ 450 AA; 50088 MW; 9DD5C0F7C129A8C1 CRC64;
 Best Local Similarity 70.8%; Score 34; DB 16; Length 450;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 MPLNAAVK 9
 Db 409 MPMLNAAVK 416

RESULT 9
 Q92FNG PRELIMINARY; PRT; 1134 AA.
 AC Q92RW9; DT 01-DEC-2001 (TREMBREL. 19, Created)
 DT 01-DEC-2001 (TREMBREL. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBREL. 21, Last annotation update)
 DE Putative sensor histidine kinase protein (EC 2.7.).
 GN R0721 OR SMCC0776
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaique V., Masuy D.,
 RA Pohl T., Portetelle D., Puebler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thébaud P., Vandebol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 Sinorhizobium meliloti strain 1021";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591787; CAC46072.1; -.
 DR InterPro; IPR00543; Y9F-like.
 DR Pfam; PF01042; UPF0076; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 154 AA; 15940 MW; 13C669F0F462BEE7 CRC64;

Query Match 68.8%; Score 33; DB 16; Length 154;
 Best Local Similarity 60.0%; Pred. No. 32;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SMPLNAAVK 10
 Db 138 SLPNAAVEI 147

RESULT 10
 Q99RUT PRELIMINARY; PRT; 459 AA.
 AC Q99RUT; DT 01-JUN-2001 (TREMBREL. 17, Created)
 DT 01-JUN-2001 (TREMBREL. 17, Last sequence update).
 DT 01-JUN-2002 (TREMBREL. 21, Last annotation update)
 DE Hypothetical protein SAV2326.
 GN SAV236 OR SA2117.

OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315);
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcaceae.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES S. aureus (strain Mu50), and S. aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11181846;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Uti Y., Takashashi N. K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Tabuchi J., Shiba T.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RT Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of meticillin-resistant Staphylococcus
 aureus";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003365; BAB58486.1; -.
 DR InterPro; IPR004770; Antiport_phac.
 DR Pfam; PF03553; Na_H_Antiporfer; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 459 AA; 48951 MW; 7342E143DDDE31D CRC64;

Query Match 68.8%; Score 33; DB 16; Length 559;
 Best Local Similarity 75.0%; Pred. No. 94;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SMPLNAAV 8
 Db 56 SMPLNAAV 63

RESULT 12
 Q90Z15 PRELIMINARY; PRT; 487 AA.
 RN [1]

SQ SEQUENCE 154 AA; 15809 MW; EB4DACAEB319FEB239 CRC64;
Query Match 66.7%; Score 32; DB 16; Length 154;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMPNRAAVM 10
|:||||| .:::
Db 138 SLPUNAPVEV 147

Search completed: March 25, 2003, 08:21:33
Job time : 26.1212 secs